

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: October 10, 2006, 14:34:58 : Search time 10518 Seconds
(without alignments)
6283.482 Million cell updates/sec

Title: US-10-623-813-85

Perfect score: 3758

Sequence: 1 MNAAVENQASGSEQVEAS.....MSVPVIRHLPAFLKEYFACV 689

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h
-O=/abs/ABSSWEB.spool/US10623813/runat_10102006_143513_7128/app.query.fasta_1
-DB=GenEmbl -GWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pcpt -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs905h
-USER=US10623813 @CGN_1_1_6706 @runat_10102006_143513_7128 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_scs:*
8: gb_sy:*
9: gb_un:*
10: gb_vl:*
11: gb_ov:*
12: gb_htg:*
13: gb_in:*
14: gb_om:*
15: gb_da:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3758	100.0	2318	6	AF480164 Mus muscu
2	3758	100.0	4094	6	BC007466 Mus muscu
3	3758	100.0	4192	6	AF068625 Mus muscu

4	3716	98.9	2597	14	AY271299	AY271299 Bos tauru
5	3716	98.9	2798	14	AY271298	AY271298 Bos tauru
6	3714	98.8	2371	5	AF480163	AF480163 Homo sapi
7	3714	98.8	3005	2	CS050243	CS050243 Sequence
8	3714	98.8	3005	5	AF067972	AF067972 Homo sapi
9	3714	98.8	3604	2	CS050242	CS050242 Sequence
10	3714	98.8	4258	5	AF311856	AF311856 Homo sapi
11	3714	98.8	4294	2	DD171352	DD171352 Marker fo
12	3714	98.8	4294	5	BC043617	BC043617 Homo sapi
13	3714	98.8	4395	2	CS050241	CS050241 Sequence
14	3697	98.4	2841	2	C0720469	C0720469 Sequence
15	3653	97.2	4476	5	AB208833	AB208833 Homo sapi
16	3586.5	95.4	3014	11	AB214886	AB214886 Gallus ga
17	3249	86.5	264380	12	AC120824	AC120824 Rattus no
18	3115.5	82.9	253474	12	AC095441	AC095441 Rattus no
19	3093	82.3	3377	11	AB196917	AB196917 Danto rer
20	3036.5	80.8	235547	12	AC112586	AC112586 Rattus no
21	2963	78.8	2933	11	AB196919	AB196919 Danto rer
22	2477.5	65.9	344787	12	AC112040	AC112040 Rattus no
23	2412	64.2	2191	5	AK025230	AK025230 Homo sapi
24	2288.5	60.9	3544	11	AB196915	AB196915 Danto rer
25	2191.5	58.3	168651	5	AC009474	AC009474 Homo sapi
26	2167.5	57.7	4267	5	AF176228	AF176228 Homo sapi
27	2164.5	57.6	4135	6	AF068627	AF068627 Mus muscu
28	2164.5	57.6	4163	6	AF151974	AF151974 Mus muscu
29	2164.5	57.6	4278	6	AF151970	AF151970 Mus muscu
30	2159.5	57.5	4043	2	C0726303	C0726303 Sequence
31	2159.5	57.5	4145	5	AF156488	AF156488 Homo sapi
32	2159.5	57.5	4195	6	AF068626	AF068626 Mus muscu
33	2159.5	57.5	4223	6	AF151973	AF151973 Mus muscu
34	2159.5	57.5	4335	2	CO981515	CO981515 Sequence
35	2159.5	57.5	4335	5	AF331857	AF331857 Homo sapi
36	2159.5	57.5	4338	6	AF151969	AF151969 Mus muscu
37	2149	57.2	3630	6	BC105922	BC105922 Mus muscu
38	2148	57.2	4369	5	D0321787	D0321787 Homo sapi
39	2056.5	53.4	2491	14	AY244710	AY244710 Bos tauru
40	2007.5	54.7	6248	11	AB196918	AB196918 Danto rer
41	1943	51.7	2421	14	AY244712	AY244712 Bos tauru
42	1892	50.3	3946	6	AF068628	AF068628 Mus muscu
43	1892	50.3	3974	6	AF151976	AF151976 Mus muscu
44	1892	50.3	4089	6	AF151971	AF151971 Mus muscu
45	1887	50.2	4006	6	AY078427	AY078427 Mus muscu

ALIGNMENTS

RESULT 1	AF480164	2318 bp	mRNA	linear	ROD 14-OCT-2002
LOCUS	AF480164				
DEFINITION	Mus musculus DNA cytosine methyltransferase 3a2 (Dnmt3a2) mRNA,				
ACCESSION	AF480164.1	GI:23954439			
VERSION					
KEYWORDS					
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Mus.				
AUTHORS	Chen,T., Ueda,Y., Xie,S. and Li,E.				
TITLE	A Novel Dnmt3a Isoform Produced from an Alternative Promoter Localizes to Euchromatin and Its Expression Correlates with Active de Novo Methylation				
JOURNAL	J. Biol. Chem. 277 (41), 38746-38754 (2002)				
REFERENCE	2 (bases 1 to 2318)				
AUTHORS	Chen,T. and Li,E.				
TITLE	Direct Submission				
JOURNAL	Submitted (31-JUN-2002) Cardiovascular Research Center, Massachusetts General Hospital, 149 13th Street, Charlestown, MA 02129, USA				
FEATURES	Location/Qualifiers				

```

source
1. .2318
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129S4/SvJae"
/db_xref="taxon:10090"
/chromosome="12"
/map="12A2-A3"
1. .2318
/gene="Dnmt3a2"
148. .2217
/gene="Dnmt3a2"
/functions="de novo DNA cytosine methylation"
/codon_start=1
/product="DNA cytosine methyltransferase 3A2"
/protein_id="AA040038.1"
/db_xref="GI:23954440"
/translation="MNAVFENQASGSKVFEASPPAVQOPTDPAAPTVAATPEPVG
DAGDKATKAAUDEPEYEDGRGIGELVWGKLRGFSWTPRGATVSWMTGRRAAGT
RWVMFSGDKFSVCVEKMLPSSFCSAHQATYNQPMYRAATIEVLQVASSRAGKL
FPACHDSDESDSGKAVEONKQMIEWALGFGQSPGKGLPEPEEKNPKEYVTDMV
EEBAAYAPPAPPAKPKSTTEKPKYKEIIDERTERLYEVAKCRNIEDICISGS
LNVTEHPLFPGMCQNCNCFLECAVOYDDGYOSCTICCGREVLMGNNCCRC
FVEECVDLLVFGGAQQA IKEDPMNMGHGHTYGLARRDMPSRMQFPANNHDO
EDDPKVTVPVPAEKRPTRVSLDPGATGLLVKDIQVDRYIYASPCDSITVG
MVRHOGKIMYVDKRSVTKHIOEMGFPLVIGSGPCNDLSIYNPARKLYGTGRLP
FEFFRLHDPAREGDDRFVWLFEENVAMGVSDKDISRFLESNPMVIMDAEVSAAH
RARFVGNLPGMNRPLASTVNDKLEQECLEHGRIAKFSKVTITITRSNIOGDOH
FVPMNEKEDILMCTEMERVFPGPVHYTDVSNMSRLARQLRGRSWSVPIRLFAPL
KEYFACV"

```

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Mismatches:	Best Local Similarity:	Percent Similarity:	Query Match:
0	2318	3758.00	689	0	0	100.0%	100.0%	100.0%
6	0	0	0	0	0	100.0%	100.0%	100.0%

US-10-623-813-85 (1-689) x AF480164 (1-2318)

```

OY 1 MetAsnAlValAlGluGluAsnGlnAlaSerGlyGluSerGlnLysValAlGluGluAlaSer 20
DB 148 ATGAATGCTGCTGGAAGAACCAAGGCTCTGAGAGCTTCAGAAAGGTGAGGAGGCGAC 207
OY 21 ProProAlaValAlGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
DB 208 CCTCTGCTGCTGACGAGCCACGACCTGCTTCTCCGACTGTGCGACACCCCTGAG 267
OY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
DB 268 CCAGTAGGAGGGGATCTGGGGACAAGAAATGCTACCAAGCAAGCCAGTAGAGCTGAG 327
OY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80
DB 328 TATGAGAGATGGCGGGGCTTTGGCATTGAGAGCTGCTGGGGGAACTTCGGGGCTTC 387
OY 81 SerTyrTrpProGlyArgIleValSerTyrTrpMetThrGlyArgSerArgAlaAlaGlu 100
DB 388 TCTGTGTGGCCAGCCGCAATTGTCTGTGTGATGACAGGCCGAGCCGAGCAGCTGAA 447
OY 101 GlyThrArgTyrValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
DB 448 GGCACCTCGCTGGCTCATGTGTGTCGAGATGCGCAAGTTCTCACTGTGTGTGTGAGAG 507
OY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
DB 508 CTCATCCCGCTGAGCTCTTCTCAGTGCACTCCACCAAGGCGACCTCAACAACAAGCAGCC 567
OY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
DB 568 ATGTACCGCAAGCCATCTACGAAGTCTCTCAGGTGCGCAGCAGCGCTGCGGGAAGCTG 627

```

```

OY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysValValAlGluLysAsn 180
DB 628 TTTCCAGCTTCCCTGACAGATGATGAAGTACAGTGCAAGGCTGTGGAAGTGCAGAAC 687
OY 181 LysGlnMetIleGluTyrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
DB 668 AAGCAGATGATTTGAATGGGCCCTGTGGTGGCTTCCAGCCCTGGGGCTCAAGGGCTGAG 747
OY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTyrAlaGluPro 220
DB 748 CCACCAAGAAAGAAAGAAATCTTACCAAGGAAGTTTACCCGACATGTGGTGAAGCTT 807
OY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
DB 808 GAAGGAGCTGCTTACGCCCAACCCACAGCCAGCAAGAAACCAAGAAAGCACAACAGAG 867
OY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
DB 868 AAACCTAAGTCAAGAGATCATTTATGAGGCAACAAAGGAGCGGCTGTGTATGAGGTG 927
OY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
DB 928 CGCCAGAAAGTCAGAAACATCGAGACATTTGTATCTCATGTGGAGGCTCAATGTCAAC 987
OY 281 LeuGluHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
DB 988 CTGGAGCAACCACTTTATTTGAGAGCATGTGCCAAGCTTGAAGACTGCTTCTTGAG 1047
OY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
DB 1048 TGTGCTTACAGTATGACGAGATGGGTACAGTCTCTATGTGACCATGCTGCTGGGGGG 1107
OY 321 ArgGluValLeuMetCysGlyLysAsnAsnAsnCysCysArgCysPheCysValAlGluCysVal 340
DB 1108 CGTGAAGGTCTCATGTGTGGGAACAACATGCTGCAAGTGTCTTTGTGTGAGAGTGTG 1167
OY 341 AspLeuLeuValAlGlyProGlyValAlaIleAlaIleLysGluAspProTyrAsnCys 360
DB 1168 GATCTCTGTGTGGGCCAGAGAGCTGCTCAAGGCACTTGAAGAAAGCCCTGGAACCTGC 1227
OY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGluAspTyrProSer 380
DB 1228 TACATGTGGGGCATTAGAGGACCTATGGGCTGCTGCCAAGCGGGAAGACTGCGCTTCT 1287
OY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
DB 1288 CGACTCCAGATGTTCTTTGCCATTAACATGACCAAGAAATTTGACCCCAAGGTTTAC 1347
OY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
DB 1348 CCACCTGTGCGAGCTGAGAGAGAGAGAGCCATCCGCTGCTCTCTTGAATGGATT 1407
OY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
DB 1408 GCTACAGGGCTCTGTGCTGTAAGGACCTGGGCATCAAGTGAACGCTACATTCCTCC 1467
OY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
DB 1468 GAGGTGTGTGAGGACTCCATCACTCGGTGGCAGATGGCGGCAAGGAAAGTCAATGAC 1527
OY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrProGlyProPheAspLeu 480
DB 1528 GTGGGGAGCTCCGAGCTGCACAGAGAGAGCCATCCAGAGATGGGGCCCATTCACCTG 1587
OY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
DB 1588 GTGATTGGAGGAGCTCCCTGCAATGACCTCTCATTTGCAACCTGCGCGCAAGGACTT 1647
OY 501 TyrGluGlyThrGlyArgLeuPheGlnPheTyrArgGluLeuHisAspAlaArgPro 520
DB 1648 TATGAGGGTACTGGCGGCTCTTCTTGAATTCACCGCTCTCAATGATGCGGGGCC 1707
OY 521 LysGluGlyAspAspArgProPhePheTyrPheGluAsnValValAlaMetGlyVal 540

```


US-10-623-813-85 (1-689) x BC007466 (1-4094)

OY 1 MetAsnAlaValGlnGlnuSnglnAlaSerGlyGlnLysValGlnGlnAlaSer 20
 DB 902 ATGAAATCGCTGGAGAGAACCGAGCCCTCGAGAGATCTCAGAAAGTTGAGGAGGCCAC 961
 OY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 DB 962 CCTCCGCTGTCAGAGACCGACCGGACCTGCTTCCACATGTCGACACCCCTGAG 1021
 OY 41 ProValGlnGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 DB 1022 CCAATGAGAGAGGAGATCTGGAGACAGAAATGCTACCAAGACCCGACCAATGAGCTGAG 1081
 OY 61 TyrGlnAspGlyArgGlyPheGlyIleGlyGlnLeuValTyrGlyLysLeuArgGlyPhe 80
 DB 1082 TATGAGGATGGCCGGGGCTTTGGCATTTGAGAGCTGTGTGGGGGAACTTCGGGGCTTC 1141
 OY 81 SerTyrTrpProGlyArgGlyValSerTyrTrpMetThrGlyArgSerArgAlaAlaGlu 100
 DB 1142 TCCTGTGGCCAGCGCCGAATGTGTCTGTGGATGACAGCGCGGAGCCGAGAGCTGAA 1201
 OY 101 GlyThrArgTyrValMetThrPheGlyAspGlyLysPheSerValValCysValGlnLys 120
 DB 1202 GGCACCTCGTGGGTGATGTGGTTGAGAGATGGCAATCTCAGTGTGTGTGGAGAG 1261
 OY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 DB 1262 CTCAATGCGCTGAGCTCTTCTTGCACTGCAATTCACCAAGCCCAACCAAGCAGCC 1321
 OY 141 MetTyrArgLysAlaIleTyrGlnValIleGlnValAlaSerSerArgAlaGlyLysLeu 160
 DB 1322 ATGTACCGCAAGACCAATCTACGAAGTCTTCAGAGTGGCAAGCCGTCCGGAAAGCTG 1381
 OY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGlnValGlnAsn 180
 DB 1382 TTTCACAGCTTGGCATACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1441
 OY 181 LysGlnMetIleGlnTyrAlaLeuGlyIleGlnIleProSerGlyProLysGlyLeuGlu 200
 DB 1442 AAGCAGATGATTAATGAGGCGCCCTCGGTGCTTCACGCCCTCGGTCTTAAGGCGCTGAA 1501
 OY 201 ProProGlnGlnGlnLysAsnProTyrLysGlnValTyrThrAspMetThrValGlnPro 220
 DB 1502 CCACCAAGAAAGAGAGAAATCTTACAGGAAGTTTACACGCAATGTGGGAGCCCT 1561
 OY 221 GluAlaAlaAlaTyrAlaProProProAlaLysLysProArgLysSerThrThrglu 240
 DB 1562 GAAAGAGCTGCTTACGCCCAACCCCAACCAAGCAAGCAAGCAAGCAAGCAAG 1621
 OY 241 LysProLysValLysGlnIleIleAspGlnArgThrArgGlnArgLeuValTyrGlnVal 260
 DB 1622 AAACCTTAAGGTCAAGAGATCATTTGATGAGCGCAAGGAGGCGGTGGTGTAGAGTG 1681
 OY 261 ArgGlnLysCysArgAsnIleGlnAspIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 1682 CCGCAAGATGCAAGAAATCGAGGACATTTGATTCATGTGGGAGCCCTCAATGTCAAC 1741
 OY 281 LeuGlnHisProLeuPheIleGlyIleGlyIleCysGlnAsnLysLysAsnLysPheLeuGlu 300
 DB 1742 CTGGAGACCCCACTTCTTATTTGAGAGCATGTGCCAAGACTGTAAAGACTCTTCTGGAG 1801
 OY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
 DB 1802 TGTGCTTACAGATGACAGATGAGTGGGTACAGTCTTATTTGACCATCTGCTGTGGGG 1861
 OY 321 ArgGlnValLeuMetCysGlyLysAsnAsnCysCysArgCysPheCysValGlnCysVal 340
 DB 1862 CGTGAAGTCTCATGTGTGGGAAACAACACTGCTGAGGCTTTGTGTGCGGTGTG 1921
 OY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGlnLysProThrAsnCys 360

DB 1922 GATCTCTGTGGGGCCAGAGACTGTCTCAGGACGCCATTAAAGAAAGACCCCTGAACTGC 1981
 OY 361 TyrMetCysGlyHisLysGlyTyrTyrGlyLeuLeuAlaArgArgGlnAspTrpProSer 380
 DB 1982 TACATGTCCGGGCAATMAAGGCACTTATGGCTGTGCGCAAGACGGGAAGACTGGCTTCT 2041
 OY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGlnLysPheAspProProLysValTyr 400
 DB 2042 CCACTCCAGATGTTCTTTGGCAATTAACATGACAGAAATTTGAACCCCAAGTTTAC 2101
 OY 401 ProProValProAlaGlnLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 DB 2102 CCACCTGTGCCAGTGTGAGAAAGAAAGCAATCCCGTGTCTGTCTTGTATGGGAT 2161
 OY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAlaAspArgTyrIleAsp 440
 DB 2162 GCTACAGGCTCTCTGTGTCTGAAGGACCTGGGACATCAAGTGGACCGCTACATTCCTC 2221
 OY 441 GluValCysGlnAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 DB 2222 GAGTGTGTGAGACTTCATACGCTGGGCAATGTGCGGACCAAGGAAAGATCATGTAC 2281
 OY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGlyTyrGlyProPheAspLeu 480
 DB 2282 GTGGGAGAGTCCGACGCTCACAGAAAGCATATCCAGAGTGGGGCCCATTCGACCTG 2341
 OY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2342 GTGATTTGAGGAGGACCTCTGCATGACCTTCATATGTCACCCCTGCGCAAGGACTT 2401
 OY 501 TyrGlnGlyThrGlyArgLeuPhePheGlnPheTyrArgLeuLeuHisAspAlaArgPro 520
 DB 2402 TATGAGGTACTGGCGGCTCTTCTTGAATCTACCGGCTCTGCATGATCGCGGGCC 2461
 OY 521 LysGlnGlyAspAspArgProPhePheTyrLeuPheGlnAsnValValAlaMetGlyVal 540
 DB 2462 AAGAGGGAGATGATCGCCCTTCTTGTGGCTTGTGAAGTGTGTGGCCATGTGGCGTT 2521
 OY 541 SerAspLysArgAspIleSerArgPheLeuGlnLysSerAsnProValMetIleAspAlaLys 560
 DB 2522 AGTGACAAAGGAGACATCTCGGATTTCTTGAGTCTAACCCCGTATGATTAACCCCAA 2581
 OY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArg 580
 DB 2582 GAAAGTGTCTGCGACACAGGCGCGTTACTTCTGGGGTAACTTCTGTGCATGAACAG 2641
 OY 581 ProLeuAlaSerThrValAsnAspLysLeuGlnLeuGlnCysLeuGlnHisGlyArg 600
 DB 2642 CTTTGGCATCCACGTGATGATGATGAGTGTGAGCTGCAAGATGTCTGAGACAGCGACA 2701
 OY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
 DB 2702 ATAGCCAAAGTTCAAGAAAGTGAAGACATTAACACAGGCTCAAACTCTTAAGACAGGGC 2761
 OY 621 LysAspGlnHisPheProValPheMetAsnGlnLysGlnAspIleLeuTyrCysThrglu 640
 DB 2762 AAAGACACAGATTTCCCGCTCTTCAATGAACAGAAAGAGAGACATCTGTGGTGA 2821
 OY 641 MetGlnArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 DB 2822 ATGAAAGGAGTGTGTGGCTTCCCGTCACTACAGACAGTCTCAACATGTAGAGCGCTTG 2881
 OY 661 AlaArgGlnArgLeuLeuGlnValArgSerTrpSerValProValIleArgHisLeuPheAla 680
 DB 2882 GCGAGGCAAGACTGTGGGCGGATCGTGAAGCGTGTGCGGTATCCGCCACTTTCGCT 2941
 OY 681 ProLeuLysGlnTyrPheAlaCysVal 689
 DB 2942 CCGTGAAGGAATATTTGCTGTGTG 2968

RESULT 3
 AF068625
 LOCUS AF068625 4192 bp mRNA linear ROD 06-DEC-1999

DEFINITION	Mus musculus DNA cytosine-5 methyltransferase 3A (Dnmt3a) mRNA,
ACCESSION	complete cds.
VERSION	AF068625
KEYWORDS	AF068625.2 GI:6449467
SOURCE	
ORGANISM	Mus musculus (house mouse)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Mus.
AUTHORS	Okanou, M., Xie, S. and Li, E.
TITLE	Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases
JOURNAL	Nat. Genet. 19 (3), 219-220 (1998)
PUBMED	9662389
REFERENCE	2 (bases 1 to 4192)
AUTHORS	Xie, S., Okanou, M. and Li, E.
TITLE	Direct Submission
JOURNAL	Submitted (28-MAY-1998) CVRC, Mass. Gen. Hospital, 149 13th Street, Charlestown, MA 02129, USA
REFERENCE	3 (bases 1 to 4192)
AUTHORS	Okanou, M., Chijiwa, T., Sasaki, H. and Li, E.
TITLE	Direct Submission
JOURNAL	Submitted (04-NOV-1999) CVRC, Mass. Gen. Hospital, 149 13th Street, Charlestown, MA 02129, USA
REMARK	Sequence update by submitter
COMMENT	On Nov 18, 1999 this sequence version replaced gi:3327977.
FEATURES	location/Qualifiers
source	1. .4192
	/organism="Mus musculus"
	/mol_type="mRNA"
	/db_xref="taxon:10090"
	/chromosome="12"
	/map="4.0 CM"
gene	1. .4192
	/gene="Dnmt3a"
	1. .4192
	/gene="Dnmt3a"
	/note="similar to EST sequences deposited in GenBank
	Accession Numbers AA052791, AA111043, AA154890, AA240794,
	AA756653, W58898, W59299, W91664, and W91665"
	217. .2943
CDS	1. .2943
	/gene="Dnmt3a"
	/function="de novo DNA methylation"
	/note="contains Cys-rich region"
	/codon_start=1
	/product="DNA cytosine-5 methyltransferase 3A"
	/protein_id="AAC40177.2"
	/db_xref="gi:6449468"
	/translation="MPSGPDTSSTSSLEREDDRKEGEEOENRKEKEKEEOPSPATARK
	VGRGKRKHPPVESDTRKDPAYTKSPAAOAGSPDLIPNDLEKRSPOGEESG
	PAAQKKGAPAECEGTEPPRAVAENWCCVTEGRGASGEDEKOKNTIEMKME
	GSRGRLGGLNBSLRLRPPRLITFOAGDYYISIKRRDMLARKREARKKAYIA
	EVNAVEENQAGESQKEASPPPAVQPTDASPLVATTPFVGDDGDKATRAAD
	EPMEEDGFGFGLVWKLKGFSPWPGRIIVSMWTKSRPAEGTRVMMWGDSKFS
	VCVEKLMLPSFSCAFHOATYKQPMYKAIYELVQVSAAGHLFPACHSDSDSGS
	KAIVENKQMIEMALGFGPSPGPIEBPEEKPNPYKEVTVDMVPEPAAAYAPP
	KPKPKSTTEKPKKEIIDERTRELVYVRQKCNIDICISGSLVTEHLPFIIGG
	MCQKCNFLECAIYDDDDGYOSTTCTCGGAEVLMGNNNCRCFCYECVDLVVGG
	AAQAKIKEDPNNCYMCGHKGTGYGLRRREDWPSRLQMFANHQDEPPKVIYPPV
	EKRRIYVLSLFDGIATGLVLKDLGIQVDYVIASEVEDSLTVGMVHQQIKMYVD
	VRSTQKHIOEMGPDIVIGSPCNLDSLVNVPARKIGEGTGLFFEYRLLHARPP
	EGDGRPFMLFENVVAVGVSKRDISRLFENPNVLDKEVSAAHARAYFPGNPGMM
	RPLASRVNDKLELQCELEHGRILAFSKRYRTITTSNSITKQKDOHPVNMKEDILM
	CTLERVFVGPVHTVDVSNMSRLARQLRLGRSWSVPVIRHLPALKEFYACV"

Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0
US-10-623-813-85 (1-689) x AF068625 (1-4192)			
QY	1 Metrsna1aValGluGluAsnGlna1aSerG1yGluSerGlnyValGluGlna1aSer	20	
Db	874 ATGAATCTGTGGAGAGAAACCAAGGCTCTGGAGAGCTCCAGAGGTGGAGAGCCAGC	933	
QY	21 ProProAlaValGluGlnProThrAspProAlaSerProThrVala1aThrThrProGlu	40	
Db	934 CCTCTGTGTGGCGAGCCCAACCGACCCTGCTTCCGACTGTGGCCACACCCTCGAG	993	
QY	41 ProValG1yG1yAspAlaG1yAspLyAspAla1aThrLyAs1a1a1aAspAspG1uProG1u	60	
Db	994 CCAATTAGAGGGAGTGTGGGGACAAAGATCTACAAAGACGCCGACGATGAGCCTGAG	1053	
QY	61 TyG1uAspG1yArG1yPheG1y1eG1yG1uDeuValTPrG1yLySLeuArG1yPhe	80	
Db	1054 TATAGAGATGGCCGGGCTTTGGCATTTGGAGAGCTGTGGTGGGAAACTTCGGGGCTTC	1113	
QY	81 SerTTrPPrProG1yArG1yAlaSerTTrPTrMetThrG1yArGSerArGAla1aGlu	100	
Db	1114 TCCGTGTGGCCAGGCCCAATTTGTCTTTGTGGATGACAGGCCGAGCCGAGCGCTGAA	1173	
QY	101 G1yThzArGrPvAlaMetTrPheG1yAspG1yLySPheserVala1CyValaG1uLyS	120	
Db	1174 GGCACGTGCGGGTCATAGTGGTTGGAGATGGCAAGTTCTCAGTGTGTGTGTGGAGAG	1233	
QY	121 LeuMetProLeuSerSerPheCySerAlaPheHisGlna1aThrTyAsnLySglnPro	140	
Db	1234 CTCATGCCGCTGAGCTCTTCTGTGAGATGCAATTCACCAAGGCCACTCAACAAGACAGCC	1293	
QY	141 MetTyArG1ySalal1eTyG1uValleuGlnVala1aSerSerArG1aG1yLyLeu	160	
Db	1294 ATGTACCGCAAGCCATCTACGAAAGTCTCCAGGTGGCCAGCACCGTCCGGGAACTG	1353	
QY	161 PheProAlaCySh1aAspSerAspG1uSerAspSerG1yLySalValaG1uValG1uAsn	180	
Db	1354 TTTCACCTTGGCCATGACAGTGAAGAAATGACAGTGGCAAGCTGTGGAAATGCAGAAC	1413	
QY	181 LysGlnMet1eG1uTPrAlaLeuG1yG1yPheGlnProSerG1yProLySg1yLeuGlu	200	
Db	1414 AAGAGATGATGTAATGGGCCCTCGGTGGCTTCAGCCCTCGGGTCTTAAAGGCGCTGGAG	1473	
QY	201 ProProGluGluG1uLySAsnProTyLySg1uValTyThrAspMetTrPvalGluPro	220	
Db	1474 CCACACAGAGAGAGAAATCTCTTAAAGAACTTAAACCCGACATGTGGTGGAGCTT	1533	
QY	221 GluAla1a1a1aTyAlaProProProProAlaLyS1ySProArG1ySserThrThGlu	240	
Db	1534 GAACACAGTGTGGCCCAACCCCAACCCCAAGCAAGAAACCCAGAAAGACACAAACAG	1593	
QY	241 LysProLySValLySg1u1e1eAspG1uArGThrArG1uArGLeuValTyG1uVal	260	
Db	1594 AAACCTTAAGTCAGAGATATATTAGAGAGGACCAAGGAGCGCGTGTATAGAGGTG	1653	
QY	261 ArgGlnLyScyArGAsn1eG1uAsp1eCyS1leSerCySg1ySer1eAsnValThr	280	
Db	1654 CGCCAGAAAGTCAGAAACATCGAGGACATTTGTATCTCATGTGGAGCCTTCAAGTCAAC	1713	
QY	281 LeuGln1ySProLeuPhe1eG1yG1yMetCySg1uAsnCyS1ySAsnCySPhaLeuGlu	300	
Db	1714 CTGAGAGCACCCACTTTCATTGAGAGCGATGCGCAAGACTGTAAAGACTGCTTTGGAG	1773	
QY	301 CysAlaTyG1uTyArAspAspAspG1yTyGlnSerTyCySThr1eCyScySg1yG1y	320	
Db	1774 TGTCTTAAACAGTATAGACAGATGGGTACAGTCTATTGCAACATCTGTGTGGGGGG	1833	
QY	321 ArgGluValLeuMetCySg1yAsnAsnAsnCyScySarGcySPhCySValG1uCyVal	340	
Db	1834 CGTAAAGTGCTCATGTGTGGGAACAACAATCGTGCAGAGTCTTTTGTGTGAATGTGTG	1893	

[illegible]

	AY271299	2597 bp	mRNA	linear	MAM 05-SEP-2003
LOCUS					
DEFINITION	Bos taurus DNA methyltransferase 3a isoform 4 (DNMT3A) mRNA,				
ACCESSION	AY271299				
VERSION	AY271299.1	GI:32350982			
KEYWORDS					
SOURCE					
ORGANISM	Bos taurus (cattle)				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.				
REFERENCE	1 (bases 1 to 2597) Golding,M.C. and Westhusin,M.E. Analysis of DNA (cytosine 5) Methyltransferase mRNA Sequence and Expression in Bovine Preimplantation Embryos, Fetal and Adult Tissues				
JOURNAL	Gene Expr. Patterns 3 (5), 551-558 (2003)				
REFERENCE	2 (bases 1 to 2597) Golding,M.C. and Westhusin,M.E. Direct Submission Submitted (08-APR-2003) Veterinary Physiology, Texas A&M - CVM, 500C University Dr. West. College Station, TX 77843-4466, USA				
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source	1..2597 /organism="Bos taurus" /mol_type="mRNA" /db_xref="taxon:9913" <1..2597 /gene="DNMT3A" 1..2467 /gene="DNMT3A" /note="DNMT3a4; DNA (cytosine 5) methyltransferase" /codon_start=2 /product="DNA methyltransferase 3a isoform 4" /protein_id="AAP5902.1" /db_xref="GI:32350983"				
gene	/translation="EGGEQEBARAKKERROESTTAAKYVGRPKRKHPVESSTPPKKK EAVTSLSMAODSGPSBLIPNGLEKSEFOPPEGGPAGOGKGAPREGATPTTP EASVAVENGCCTPEKPELOKKAERAKYIAWNAVEENOGSTESORVEASFPAVOQFP DPAEPVTATPEPVGADAGCAKNATKADDEBEYEDGNFGIGELMWGLRFSSWPMPG IVSWMTGRSJAAGETRWMMFMGDKEFSVVCVELAMPDSFCSAFHQATVKOMPYRA ALIVELOVASBAKGLPFMGCDSPDETAKAEYONKOMIEMALGFOSGPKELESRP EEENPKRYEVYTDMWVRPEAAAYAPPAPAKPRSTTEKPEVKEIIIBERTERYLVE ROKRNIIEDICISCSLANVLIEHLPIFGCMQNCKNPLECAVGYDDDGOSTYCIC GGRVLMCGNNCCRCFCVECDVLLVGGAQAQAIKEDPNKCYCGHKGTGLRRR DWPRLOMFANNDHOFDEDPKVYPVPVAPERKPIRVALSIFDLATGLLVKIDGITQ DRYTAISEVEDSIITGVGWROHGKIMYDVASTVQKHIOEWGPDLYVGSBCNDLS VNPARKGIECTGTLFFEFYLLHDHARDKEGDDPPFLFNENVVLGVSDKRLSR ESNVMDIAKEVSAAHRARYPMGLPCMNPLSLTVDKLELOCLEHGRJAKSKSVN TTTTRSNIRKQDOHFVFNMKEKDILMCTEMERVGCFPHYTDVSMNSLAQRLLT GRMSVPIRHLFAPLKYEFACV"				
CDS					
ORIGIN					
Alignment Scores:					
Pred. No.:	0	Length:	2597		
Score:	3716.00	Matches:	680		
Percent Similarity:	99.1%	Conservative:	3		
Best Local Similarity:	98.7%	Mismatch:	6		
Query Match:	98.9%	Gaps:	0		
DB:	14				
US-10-623-813-85 (1-689) x AY271299 (1-2597)					
Oy	1 MetanalaValAlGUlGuLaenGlnAlaSerGlyGuSerGlnLyVaAlGUlGuLaLas	20			
Dd	398 ATGATGCTGTGTGAAGAAACCAGGGGCCCCAGGTCAGAAAGTGACAGAGGCCCAT	457			
Oy	21 PropoAlaValAlGnGInPProThAspProAlaSerProTrnValAlaThThrProGlu	40			
Dd	458 CTCTCCGCTGTGCACGACCCACGACCCCTCGCTCCCCACAGTGGCCACACGCGCTAG	517			
Oy	41 ProValGlyGlyAspPalagLyAspLyAsnAlaThrLyAlaAlaAspAspGluProGlu	60			

```

Db      ||||| 518 CCGCTGGGCGCCGATGCGGGGCAAGAAATGCCAACCAGCTGACATGAAACCGGAG 577
Qy      ||||| 61 TGTGAASPGIYARGIYPHEGIYIEGIYGIULEUVALTRPGIYLYSEUARGIYPHE 80
Db      ||||| 578 TACGAGACGGCGGGGCTTTGGCATTTGGGGAGCTGGTGTGGGGAACCTGGCGGCTTC 637
Qy      ||||| 81 SETPTTPROGIYARGIILEVALSETPTTPMETHTGIVYRGERARGIAAIAAGIU 100
Db      ||||| 638 TCCGTGTGGCCAGGCGCGCATTTGTGTGTGATGACGGGCGGAGCCGAGCGGGA 697
Qy      ||||| 101 GIYTHARGTRVALMETTRPHEGIYASPGIYLYSPHESEVALIYCYVALGIULY 120
Db      ||||| 698 GGCACCTGTGGGTCTGTGTGTTCCGAGACGCGAACTTCTCACTGGTGTGTGGAGAG 757
Qy      ||||| 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTYRAnlysginPro 140
Db      ||||| 758 CTGATGCGCGTGAAGCTCTTCTGACAGTGTTCACACAGGCCACCTACAAACAGCAGCCC 817
Qy      ||||| 141 MetTYRARGLYRAlAlIeTYRGIUVALIeUGIIVALIAsERSEARGIAAGIYLYBLEU 160
Db      ||||| 818 ATGTACCCGACAGGCTATCTACAGAGTCTCGAGGTGGCCAGCATGAGCGGGGAAGCTG 877
Qy      ||||| 161 PHEPROAlaCYeHisAspSerAspGluSerAspSerGlyLYsAlaValGIUValGlnAsp 180
Db      ||||| 878 TTTCCAAATGTGCAATACAGCCGACGAGAGCGACACTGCGCAAGGCCCTGGAGGTGCAAGAC 937
Qy      ||||| 181 LysGlnMetIIEGIUTRPAIeUGIYIeGIYPhEgInProSerGIYProLYsGIYLeUGIU 200
Db      ||||| 938 AAACAATAATGATGAGTGGGCGCTTGAGAGGTTCCAGCCCTCTGGCCCAAGGGCGCTGAG 997
Qy      ||||| 201 PROPTRGUGIUGIULYsAsnProTYRLYRGIUVALIYTRThAspMetTRPVALGIUPro 220
Db      ||||| 998 CCCCCGAAAGAGAGAAAGAACCCCTCAAGAAATTACACAGACTGTGGGTGAACCC 1057
Qy      ||||| 221 GIUAlaAlaIaTYRAlaProProProProAlaLYsLYsProARGLYsSERTRThRGIU 240
Db      ||||| 1058 GAGGACGCTGCTATGCGCGCGCCCAACGACCAAAAGCCCCGAGAAAGCACAGAGAG 1117
Qy      ||||| 241 LysProLYsValIeSGIUIeIleAspGIUARGThRARGIULARGLEUVALIYRGIUVAL 260
Db      ||||| 1118 AAGCCTAAGGTCAAGAGATCATTTGATGAACGCAAGAGAGCGGCTGTGTGCGAGTA 1177
Qy      ||||| 261 ARGGLIYECYsARGsAnIIEGIUAspIIECYsIIEsERCYsGIYsERLeUAsnVALTr 280
Db      ||||| 1178 CGCGAAGACTGCGGAAACATCGAGGACATCTGCATCTCTTGTGGAGCCCTCAACGTACCC 1237
Qy      ||||| 281 LeuGIUHisProLeuPheIIEGIYGIYMETCYsGlnAsnCYsLYsAsnCYsPheLeUGIU 300
Db      ||||| 1238 TTGGAAACCCCTCTCTTCATCGAGGAATGTGCCAAACTGCAAGAACTGCTTCTGGAA 1297
Qy      ||||| 301 CYsAlaTYRGIUTrAspAspAspGIYTYRGIUSeTYRCYsThRIIECYsCYsGIYGIY 320
Db      ||||| 1298 TCGGCTTACAGTACATGACATGATGCTGCTACCTCACTGACCACTGTCTCGGGGGG 1357
Qy      ||||| 321 ARGGIUVALIeUeCYsGIYsAsnAsnAsnCYsCYsARGCYsPheCYsVALGIUCYsVAL 340
Db      ||||| 1358 CCGGAGGTGCTCATGTGTGGGAAACAATTGCTGAGGGGCTTTTGGCGGTGATGTGTG 1417
Qy      ||||| 341 AspLeuLeuValGIYProGIYAlaIaGlnAlaAlIeLYsGIUAspProTPAsnCYs 360
Db      ||||| 1418 GATCTCTTGTGGTGGGCGAGGGGCGCGGACGACCATCAAAAGAAACCCCTGAAACTGC 1477
Qy      ||||| 361 TYRMeCYsGIYHisLYsGIYThTYRGIYLeuLeuARGIARGIULAspTRProSer 380
Db      ||||| 1478 TACATGTGTGGGACAAAGGCACTACGGGCTGTCTGGGGGGGCGAGCGAGCTGGCTGT 1537
Qy      ||||| 381 ARGLeUGIeMeCPhePheAlaAsnAsnHisAspGlnGIUAspProProLYsVALIYTR 400
Db      ||||| 1538 CGGCTTCCAGATGTCTTCCGCAACAACATGACAGAGAAATTCATCTCCGAAAGGTTTAC 1597
Qy      ||||| 401 PROPTRVALProAlaGIUYSARGLYsProIleARGVALeUSeRLeuPheAspGIYIle 420

```

```

Db      ||||| 1598 CCACCTGTCCGACGTGAGAAAGAAAGCCATCCGGGTGTCTGTCTATTCATGTAAT 1657
Qy      ||||| 421 AlaThRGYLeuLeuValIeUlyAspLeUGIYIIEGIUValAspARGTYRIIEAlAsER 440
Db      ||||| 1658 GCTACAGGGCTTCTGTGTGTGAAGACCTGGGATTCAGGTGAGCGCTACATGCGCTCC 1717
Qy      ||||| 441 GIUValCYsGIUAspSERIIEThRVALGIYMETVALARGHisGlnGIYLYsIIEMeTYR 460
Db      ||||| 1718 GAGGTGTGAGAGACTTCTACACGTGGGCAATGTGTGGGCAACAGGAAAGATCATGTAC 1777
Qy      ||||| 461 VALGIYAspVALARGSERVALThRGInLYsIIEGIUValTRPGIYProPheAspLeu 480
Db      ||||| 1778 GTCCGGGACGTCCGACGCTTACACAGAGCATATTCAGAAATGGGGCCGCTGCAATCTG 1837
Qy      ||||| 481 VALIIEGIYLYSERProCYsAsnAspLeuSERIIEVALAspPROAlaARGLYsGIYLeU 500
Db      ||||| 1838 GTGATTTGGGGGAGTCCCTGCAATGATCTCTCATGTCAACCTTCGCGCAAGGAGCTC 1897
Qy      ||||| 501 TYRGIULYThRGYARGLeuPhePheGIUPhETYRARGLeuLeuHisAspAlaARGPro 520
Db      ||||| 1898 TACGAGGCACTGGCGGCTCTTTTGAGTCTACCGCTCTCTGCATGATGGCGGCCC 1957
Qy      ||||| 521 LysGIULYAspAspARGProPhePheTRPLeuPheGIUAsnVALAlaMeTGIYVAL 540
Db      ||||| 1958 AAGGAGGAGATGACCGGCCCTTCTGTGCTTGAATAATGTGGTGTCTGGGCGCTT 2017
Qy      ||||| 541 SerAspLYsARGAspIIEsERARGPheLeUGIUsERAsnProVALIeMetIIEAspAlaLYs 560
Db      ||||| 2018 AGTGACAAAGGAGACATCTCGCATTTCTGAGTCCAACCTGTGATGATTGATGCCAAA 2077
Qy      ||||| 561 GIUValSERAlaAlaHisARGIAlaARGTYRPhETRPGIYAsnLeuProGIYMeTAsnARG 580
Db      ||||| 2078 GAAGTCTCAGCTGCGGACAGGGCCGCTACTTCTGGGGAACTTCTGTGATGAACAGG 2137
Qy      ||||| 581 PROLeuAlAsERThRVALAsnAspLYsLeUGIULeUGIUCYsLeUGIUHisGIYARG 600
Db      ||||| 2138 CCATTTGCATCCACTGTGATGATTAAGTGAAGCTCAGAGTCAAGAGTGTGGAGCAGCGCGA 2197
Qy      ||||| 601 IIEAlaLYsPheSERLYsVALARGThRIIEThRARGSERsAnSERIIELYsGInGIY 620
Db      ||||| 2198 ATAGCCAACTTACGACAAATGAGAACCATTACTACTAGGTGCAATCCAAAGACAGGGC 2257
Qy      ||||| 621 LysAspGlnHisPheProVALPheMetAsnGIULYsAspIIELeuTRPYsThRGIU 640
Db      ||||| 2258 AAGAGCACCAATTTCCCGCTTTCATGATGAGAAAGAGACATCTTATGTGTACATGAA 2317
Qy      ||||| 641 MetGIUARGVALPheGIYpHeProVALHisTYRThAspVALSERAsnMeTSErARGLeU 660
Db      ||||| 2318 ATGAAAGGTGTGTGGCTTCCCTGTCACTATACCGAGGTCTCCAACATGAGCCGCTTG 2377
Qy      ||||| 661 AlaARGGlnARGLeuLeUGIYARGSERTRPSErVALProVALIIEARGHisLeuPheAla 680
Db      ||||| 2378 GCGAGGCAAGACACTGTGGCGGTGTGTGAGAGGTGCCGTGATCCGCCACTTTCGCT 2437
Qy      ||||| 681 PROLeuLYsGIUTrYPheAlaCYsVAL 689
Db      ||||| 2438 CCGCTGAAGGAATATTTCTGTGTG 2464

```

RESULT 5

AY271298 2798 bp mRNA linear MAM 05-SEP-2003
 LOCUS Bos taurus DNA methyltransferase 3a (DNMT3A) mRNA, partial cde.
 DEFINITION
 AY271298
 ACCESSION
 VERSION AY271298.1 GI:32350980
 KEYWORDS
 SOURCE Bos taurus (cattle)
 ORGANISM Bos taurus

REFERENCE
 AUTHORS
 TITLE
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 2798)
 Golding, M.C. and Westhusein, M.E.
 Analysis of DNA (cytosine 5) Methyltransferase mRNA Sequence and

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Expression in Bovine Preimplantation Embryos, Fetal and Adult
Tissues
Gene Expr. Patterns 3 (5), 551-558 (2003)
2 (bases 1 to 2798)
Golding, M.C. and Westhusin, M.E.
Direct Submission
Submitted (08-APR-2003) Veterinary Physiology, Texas A&M - CVM, 500
University Dr. West, College Station, TX 77843-4466, USA

FEATURES

source
1. 2798
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
<1. 2798
/gene="DNMT3A"
<1. 2668
/gene="DNMT3A"
/note="DNMT3A: DNA (cytosine 5) methyltransferase"
/codon_start=2
/product="DNA methyltransferase 3a"
/protein_id="AAP75901.1"
/translation="EGEBOEBARAKEBOEPSTTARKVGRPGRKHPVESDTPKD
/db_xref="GI:32350981"
/translation="EGEBOEBARAKEBOEPSTTARKVGRPGRKHPVESDTPKD
PATVSKLSMAODSGRSELYPNGLDKRSPQEBGSPAGQKGAAPAGEGATERP
EASRAVENGCCTKDGRCAPAEBSKEQKENTISMCEGRGLRGGLMESLRRP
MRLTFQAGDPYIISRKDEWLARKRELEKAKYIAVWNAVEVNGSTESQKZEA
SPPAVOQPDPAPEVATVPEPVGADADNATKADDEVEDGSGFGLGELWKL
RGSFWMGRIVSWMTGRSAAEGTRVWVFGDKFSVVCVELMSEFSCAFHQAT
YNNQPMYRAIVLQVASSRAGLPPMCHDSDESDTAKEVONKQIMLALGQOP
SGPKGLEPBEENPYKEVYTDWVPEAAVAPPAKPKRSTKKEPKVXIIIDER
TERRLYEPRQKRNIEDICISGSLNLTLEHPLFCGMCONKNCLECAVYDDG
YOSYCTCCGGRVLMGNCCRCCEVCVDLLVPGAAQAIKEDPMNCYCGKRG
TYGLARRDMPRLQMFANNDQEPDPKVPVPAEKRTIRVLSLDGATLIL
VKDLGIIQVDRYIASEVEDSITVGMWRHQKIMYGDVRSVQKIQEMPPDLVIG
GECNDLSIVNPARKGLYEGTGRLEFEFYLADARPKGDDPFFMLEMNVLVGS
DKRDISRLPESNPMVMDAKEVSAHAHRAVEMGNLPGNRLAETVNDKLEOCLBHG
RIAKESKVTITTRSNSIKOGKDOHPFVENEKEDILMCTEMRVRGFPVHTYDVSNM
SRLARORLGRSWSVPVIRHLRPLKEFYACV"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 2798
Score: 3716.00 Matches: 680
Percent Similarity: 99.1% Conservative: 3
Best Local Similarity: 98.7% Mismatches: 6
Query Match: 98.9% Indels: 0
DB: 14 Gaps: 0

US-10-623-813-85 (1-689) x AY271298 (1-2798)

QY 1 MetAsnaIaValGluGluAenGlnAlaSerGlyGluSerGlnLysValGluGluAlaSer 20
DB 599 ATGAATGCTGTTGAAGAAACAGAGGCTCCACCGAGCTCAGAGGCGAGGCGCAGT 658
QY 21 ProProAlaValGlnGlnProThraSProAlaSerProthraValAlaThrTrpProGlu 40
DB 659 CCTCCCGCTGTGAGCGCCACCGACCTGCGGCCGACAGTGGCCACCGCCTGAG 718
QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
DB 719 CCGGTGGGCGCCATGCGCGGAGCAAGAAATGCCACCAACGAGTGCATGAAACCGGAG 778
QY 61 TyrGluAspGlyArgGlyPheGlyTILEGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
DB 779 TACGAGACGCGCGGCTTTGGCATGGGAGCTGTGTGGGGAAACTGCGGGCTTC 838
QY 81 SerTrpTrpProGlyArgGlyLeuValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
DB 839 TCTGTGGGCGACGCGCCATGCTGTCTTGATGATGACGGGCGGAGCCGACGCGGAA 898
QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
DB 899 GGCACTGCTTGGGTCATGTGTTGAGAGACGGCAAGTTCTCAGTGTGTGTGAGAGAG 958

QY 121 LeuMetProLeuSerSerPheCysSerAlaPheIleGlnAlaThrTyrAsnLysGlnPro 140
DB 959 CTGATGCGCGTGAAGTCTCTTGTGAGTGTCTTCCACAGGCCACTTACACAGACGCC 1018
QY 141 MetTyrArgLysAlaIleTyrGluValIleuGlnValAlaSerSerArgAlaGlyLysLeu 160
DB 1019 ATGTACCGAAGGCTATCTACAGAGGCTCTGACGAGTGGCCAGACGTCAGCGGGAAGCTG 1078
QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluGlnAan 180
DB 1079 TTTCCAAATGTCATGACACGACGACGACGACGACGACGACGACGACGACGACGAC 1138
QY 181 LysGluMetIleGluTrpAlaLeuGlyIleGlyPheGlnProSerGlyProLysGlyLeuGlu 200
DB 1139 AAACAAATGATGAGTGGGCTCTGGGAGGCTTCCAGCCTTGGCCCAAGGCGCTCGAG 1198
QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
DB 1199 CCCCAGAAAGAGAGAGAGAACCCCTACAAAGAGTTACACAGACATGTGGGTTGAACCC 1258
QY 221 GluAlaAlaIleTyrAlaProProProProAlaLysLysProArgLysSerThrTrpGlu 240
DB 1259 GAGGACGCTGCTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1318
QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLysValTyrGluVal 260
DB 1319 AAGCTTAAGTCAAGAGAGATCAATTGTAACCAACAGAGAGCGCTGTCACAGGTA 1378
QY 261 ArgGluLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
DB 1379 CGCGAAGATGCGGGAATCGAGACATCTCATCTTGTGGAGCGCTCAACGTCACCC 1438
QY 281 LeuGluHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
DB 1439 TTGGAAACCCCTCTCTTCTCGAGAGATGTCACAAATCGCAAGATGCTTCCGGA 1498
QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
DB 1499 TGGCGTACCATGATCATGATCATGATGCTCATGCTTATGACATCTGCGGGGG 1558
QY 321 ArgGluValIleuMetCysGlyAsnAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
DB 1559 CGGAGAGTCTCATGTGTGGAACAACATTTGCTGAGGTGCTTTGCGGTGAATGTGTG 1618
QY 341 AspLeuLeuValGlyProGlyAlaIleGlnAlaIleLysGluAspProTrpAsnCys 360
DB 1619 GATCTCTTGTGGGCGCAGGCGCGCGCGCGCGCGCGCATCAAGAGAACCCCTGGAACTGC 1678
QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGlyLysProSer 380
DB 1679 TACATGTGCGGCGACAGAGGACCTACGCGGTGCTGCGCGGGGAGCGACTGGCGGTCT 1738
QY 381 ArgLeuGluMetPhePheAlaAsnAsnHisAspGlnGluPheAspProPolysValTyr 400
DB 1739 CGGCTCCAGATGTTCTTCCCAACAACATGACACAGAAATTCATCTCCGAAAGTTTAC 1798
QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyLe 420
DB 1799 CCACTGTCCAGCTGAGAGAGAAAGCCATCCGGGTGTGCTCTTATTCATGAGAAAT 1858
QY 421 AlaThrGlyLeuLeuValIleuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
DB 1859 GCTACAGGCTTCTGTGCTGTAAGACTTGGGCAATTCAGGTGACCGCTATCATCGCTCC 1918
QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
DB 1919 GAGGTGTGAGGACCTCATCACTGCGGATGTGTGCGGACCAAGGAAAGATCATGTAC 1978
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
DB 1979 GTGCGGAGAGCTCGCAGCGCTTACACAGAGCATATTCAGGAATGGGGCCGTTGCAATCTG 2038

QY 481 ValIleGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLeuGlyLeu 500
 Db 2039 GTGATTGGGGGCACTCCCTGCAATGATCTCCATCTGCAACCTCCCGGAGGACTC 2098
 QY 501 TTTGTTGTTThGTAArgLeuPheGluPheTyrTArgLeuLeuHisAspAlaArgPro 520
 Db 2099 TACAGGGCACTGGCCGGCTCTTCTTGAGTTCTACCGCCCTCCGATGATGCGCGGCC 2158
 QY 521 LysGluGlyAspAspArgProPhePheTyrPheGluAsnValValAlaMetGlyVal 540
 Db 2159 AAGGAGGAGATGACCCGCCCTTCTTCGGCTCTTTAGAAATGCGTCTCGGGCTT 2218
 QY 541 SerAspLysArgAspLysArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
 Db 2219 AGTCACAAGAGGACATCTCCGCGATTCTCGAGTCCAAACCTCGATGATGATGCCAAA 2278
 QY 561 GluValSerIleAlaHisArgAlaArgTyrPheTyrPheGluLeuLeuProGlyMetAsnArg 580
 Db 2279 GAAGTTCACCTGGCAACAGGCGCCGCTACTTCGGGGAACTTCCTGATGAAACAGG 2338
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGluCysLeuGluHisGlyArg 600
 Db 2339 CCATTGGCATCTCACTGCAATGATTAAGCTGAGCTGCAAGAGTCTTGGAGCCGCGCA 2398
 QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
 Db 2399 ATAGCCAAAGTTACAGCAAGTAGAGACCATTACTACTAGTTCGAACTCCATTAAGCAGGCG 2458
 QY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTyrCysThrGlu 640
 Db 2459 AAGGACCAAGATTTCCCGCTTCATGATAGAAAGAAAGAGACATCTTATGATGACAGAA 2518
 QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 Db 2519 ATGGAAAGGCTTTGGCTTCCCTGCTCACTATACCAAGCTCTCCAAACAGAGCCGCTTG 2578
 QY 661 AlaArgGlnArgLeuLeuGlyArgSerTyrSerValProValIleArgHisLeuPheAla 680
 Db 2579 GCGAGCAGAGACTGCTGGGCGCGTGTGAGCGTGCATCCGCCACCTCTTGCT 2638
 QY 681 ProLeuLysGluTyrPheAlaCysVal 689
 Db 2639 CCGCTGAAGGAATATTTGCTGTGTG 2665
 RESULT 6
 AF480163 2371 bp mRNA linear PRI 14-OCT-2002
 LOCUS Homo sapiens DNA cytosine methyltransferase 3A2 (DNMT3A2) mRNA,
 DEFINITION complete cds.
 ACCESSION AF480163
 VERSION AF480163.1 GI:23954437
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 REFERENCE 1 (bases 1 to 2371)
 AUTHORS Chen, T., Ueda, Y., Xie, S. and Li, E.
 TITLE A Novel Dnmt3a Isoform Produced from an Alternative Promoter
 Localizes to Euchromatin and Its Expression Correlates with Active
 de Novo Methylation
 JOURNAL J. Biol. Chem. 277 (41), 38746-38754 (2002)
 PUBMED 12138111
 REFERENCE 2 (bases 1 to 2371)
 AUTHORS Chen, T. and Li, E.
 TITLE Direct Submision
 JOURNAL Submitted (31-JAN-2002) Cardiovascular Research Center,
 Massachusetts General Hospital, 149 13th Street, Charlestown, MA
 02129, USA
 FEATURES
 source 1..2371
 /organism="Homo sapiens"

ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2371
 Score: 3714.00 Matches: 679
 Percent Similarity: 98.8% Conservative: 2
 Best Local Similarity: 98.5% Mismatches: 8
 Query Match: 98.8% Indels: 0
 DB: 5 Gaps: 0
 US-10-623-813-85 (1-689) x AF480163 (1-2371)
 QY 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyLysSerGlnLysValGluGluAlaSer 20
 Db 217 ATGAATGCTGTGAGAAACCAAGGGGCCCCGGGAATCTCAAGAGTGAAGAGGCCAGC 276
 QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 Db 277 CCTCGTGTGACAGAGCCCACTGACCCCGCATCCCGCATGTGCTACAGCGCTGAG 336
 QY 41 ProValGlyLysAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 Db 337 CCGTGGGGTCCGATGCTGGGAGCAAGATGCCACCAAGCAGCGCATGACGACGAG 396
 QY 61 TTTGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrrGlyLysLeuArgGlyPhe 80
 Db 397 TAGAGAGAGCGCGGGGCTTGTGCAATGGGGAGCTGTGTGGGGAACTGCGGGGCTTC 456
 QY 81 SerTrrPrrProGlyArgIleValSerTrrPrrMetThrGlyArgSerArgAlaIaIaGlu 100
 Db 457 TCCTGTGTGCGAGCGCGCATTTGTCTGTGTGTGAGAGCGGCGGAGCGGAGAGCTGAA 516
 QY 101 GlyThrArgTrrValMetTrrPheGlyAspGlyLysPheSerValAlaCysValGluLys 120
 Db 517 GGCACCGCGTGGGTCAATGCTGTGCGAGCGCAAAATCTCAGTGTGTGTGAGAG 576
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrTrasnLysGlnPro 140
 Db 577 CTGATGCGCGTGAAGCTGTTTTCAGTGGCTTCCACAGGCCACGTAACAAGCAGGCC 636
 QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 Db 637 ATGTACCGGAAGGCATCTACAGAGTCTCGACAGTGTGGCCAGAGCGCGCGGGAAGCTG 696
 QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluGlnAsn 180
 Db 697 TTCCCGGTGTGCGACAGCAGCATGAGTGAACACTGCCAAGCGCGTGAAGGTGCGAAGC 756

```

Qy 181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db 757 AAGCCCATGATGTAATGGCCCTGGGGGGCTTCCAGCCTTGCCCTTAAGAGCGCTGGAG 816
Qy 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
Db 817 CCACCAAGAGAGAGAAATCCCTCAAGAGATGTACAGGACATGTGGTGGAACT 876
Qy 221 GluAlaIleAlaTyrAlaProProProAlaLysLysProAlaGlySerThrThrGlu 240
Db 877 GAGGACAGCTGCTACGACCACTCCACACCAAAAGCCCGGAAAGACACGCCGAG 936
Qy 241 LysProLysValLysGluIleIleAspGluArgThrArgLysLeuValTyrGluVal 260
Db 937 AAGCCCAAGGTCAAGAGATTATGATGAGCGCAAGAGAGCGGTGTGTACGAGGTG 996
Qy 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db 997 CGGACAGAGTCCCGGAACATTGAGACATCTGCATCTCTGTGGAGCCCTCAATGTACC 1056
Qy 281 LeuGlnLysProLeuPheIleGlyGlyMetCysGlnAsnLysCysLysAsnLysPheLeuGlu 300
Db 1057 CTGGAACACCCCTCTGCTTGAGAGATGTGCCAAACTGCAGAACTGCTTCTGGAG 1116
Qy 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db 1117 TGTCGTAACCACTACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1176
Qy 321 ArgGluValLeuMetCysGlyLysAsnAsnLysCysCysArgCysPheCysValGluCysVal 340
Db 1177 CGTAGAGTGTCACTAGTGTGGGAAACAACTGTGAGGTCTTTGCTGTGAGTGTGTG 1236
Qy 341 AspLeuLeuValGlyProGluValAlaGlnAlaIleLysGluAspProTrpAsnLys 360
Db 1237 GACCTCTTGTTGGGGGCGGGGGGCTGCCAGGACCACTTAAGAAAGCCCTTGAACTGC 1296
Qy 361 TyrMetCysGlyIleLysGlyTyrThrTyrLeuLeuArgArgGluAspTrpProSer 380
Db 1297 TACATGTGGGGGCAAGAGGTACTACCGGCTGTGCGGGCGGCAAGAGACTGTGCGCTCC 1356
Qy 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProLysValTyr 400
Db 1357 CGGCTCAGAGTTCCTGTCTTAATPAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1416
Qy 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 1417 CCACCTGTCCACGCTGAGAAAGAGAAAGCAATCCGGGTGCTGTCTCTTGTGATGGAAATC 1476
Qy 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db 1477 GCTACAGGGGCTCTGTGTCTGAGAGACTTGGGCAATTCAGTGGAGCCGCTACATTCCTCG 1536
Qy 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db 1537 GAGGTGTGTGAGCACTCATCAAGGTGGCATGTGGCGCAACAGGGAGAAATCATGTATC 1596
Qy 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrGlyProPheAspLeu 480
Db 1597 GTCGGGAGACGTCGCGAGCTGACACAGAAAGCATTCAGAGTGGGGCCCTTGTGATGTG 1656
Qy 481 ValIleGlyLysProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db 1657 GTGATGGGGGCAATCTCCATGACCTCTCCATCTCCATCTCCATCTCCATCTCCATCTCC 1716
Qy 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db 1717 TACAGAGGACGTGGCGGCTCTTCTTGAATTCACCCCTCTCTGATGATGCGGCGCC 1776
Qy 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaIleMetGlyVal 540
Db 1777 AAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1836

```

```

Qy 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db 1837 AGTGACAGAGGAGACATCTCGCATTTCTCGATCTCAACCTGTGTGATGATGATGATGATG 1896
Qy 561 GluValSerAlaIleHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db 1897 GAGGTGTGACCTGACACAGAGGCGCCCTACTTCTGGGGTAACTTCCGGTATGAACAGG 1956
Qy 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnLysCysLeuGluHisGlyArg 600
Db 1957 CCGTGTGATCCACTGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2016
Qy 601 IleAlaLysPheSerLysValArgThrIleThrArgSerAsnSerIleLysGlnGly 620
Db 2017 ATGCAAGATTCAGCAAGAGAGACATTACTACAGAGGTCACAGATCTCAATTAAGCGCC 2076
Qy 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db 2077 AAGACCAAGATTTCTCTGCTTCTGATGATGATGATGATGATGATGATGATGATGATG 2136
Qy 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db 2137 ATGGAAGGATATTGTTGTTCCAGTCCACTATATGACGCTCTCCACATGAGCGCTTG 2196
Qy 661 AlaArgGluArgLeuGluArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db 2197 GCGAGGACGAGTGTGCTGGGCGGCTGACGATGAGAGCTGTCCATCTCCGCCACTTGTG 2256
Qy 681 ProLeuLysGluTyrPheAlaCysVal 689
Db 2257 CCGCTAAGAGATATTTTGTGTGTGTG 2283

```

```

RESULT 7
CS050243
LOCUS CS050243 3005 bp DNA linear PAT 23-MAR-2005
DEFINITION Sequence 27 from Patent WO2005021757.
ACCESSION CS050243
VERSION CS050243.1 GI:61889498
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE
1 Tomme, P.H. and van Rompaey, L.
TITL Polypeptides and polynucleotides for use as a medicament
JOURN Patent: WO 2005021757-A 27 10-MAR-2005;
Galapagos Genomics N.V. (BE)
FEATURES
source
location/Qualifiers
1..3005
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3005
Score: 3714.00 Matches: 679
Percent Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatches: 8
Query Match: 98.8% Indels: 0
DB: 2 Gaps: 0
US-10-623-813-85 (1-689) x CS050243 (1-3005)
Qy 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyLysSerGlnLysValGluGluAlaSer 20
Db 899 ATGAATCTCTGGAAGAAACAGAGGGCGCGGGGAGTCTCAAGAGGTGAGAGAGCCAGC 958
Qy 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrTrpGlu 40
Db 959 CTTCTCTGTGAGAGCACTGACCCCGCATCCCGCATCTGCTGTGCTGCTGCTGCTGCTGCTG 1018

```

QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 DB 1019 CCCGCGGGGCTCCGATGCTGGGGACAAGATCCACCAAGCGCGGATGACGAGCCAAAG 1078
 QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuAspArgGlyPhe 80
 DB 1079 TACAGAGACGGCCGGGGCTTTGGCATTGGGAGCTGCTGTGGGGAATCTCGGGGCTTC 1138
 QY 81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
 DB 1139 TCCCGGTGGCCAGGCGCGCATGTGTCTGTGGATGAGCGGGCCGAGCGGAGCACTGAA 1198
 QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
 DB 1199 GGCACCCGCTGGTCAATGTGTGGAGACGCAAAATTCAGTGTGTGTGTGAGAG 1258
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 DB 1259 CTGATTCGCGCTGAGCTCTTTTGGAGTGCCTTCCACCGGCGCAGTACACAGCAGCC 1318
 QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 DB 1319 ATGTAACGCAAGCATCTACAGAGTCTCTGACGGTGGCCACAGCCGCGGAGAGCTG 1378
 QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValAsn 180
 DB 1379 TTCCTGGTGTCCACGACGAGATGAGTACATCTCCAGCCGCTGTGAGCTCAGAAC 1438
 QY 181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProGlyLysGlu 200
 DB 1439 AAGCCCATGATTGAATGGCCCTTGGGGGCTTCCAGCTTCTGGCCCTTAAGGGCTGGAG 1498
 QY 201 ProProGluGluGlyLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
 DB 1499 CCACCAAGAGAGAGAAATCCCTACAAAGATGTACAGCAGATGTGTGTAACCT 1558
 QY 221 GluAlaAlaAlaTyrAlaProProProAlaLysLysProAlaGlySerThrThrGlu 240
 DB 1559 GAGGACCTGCTVAGGACCCCTCCACAGCAAAAGCCCCGGAAGAGACAGCGAG 1618
 QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
 DB 1619 AAGCCCAAGGTCAAGAGATTAATGATGACGCGCACAGAGCGGCTGTGTAAGAGTG 1678
 QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 1679 CGGCAAGATGCGCGAATTCATGAGACATCTGCAATCTCTGTGGAGGCTCAATGTACC 1738
 QY 281 LeuGlnHisProLeuPheIleGlyGlyMetCysGlnAsnGlyLysAsnCysPheLeuGlu 300
 DB 1739 CTGAAACACCCCTCTTCTGTGGAGATGTGCCAAACGCAAGAACTGCTTCTTGAG 1798
 QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
 DB 1799 TGTGCGTACCAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1858
 QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnGlyCysArgCysPheCysValGluCysVal 340
 DB 1859 CGTGAAGTGTCTCATGTGCGGAAACAACTGTCAGGTGCTTTTGTGAGAGTGTG 1918
 QY 341 AspLeuLeuValGlyProGlyValAlaGlnAlaAlaIleLysGluAspProTrpAsnGly 360
 DB 1919 GACCTTGTGGGGGCGGGGGCTGCGCCAGGACCATTAAGAGAGAGAGAGAGAGAGAG 1978
 QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGluAspTrpProSer 380
 DB 1979 TACATGTGGCGGCAAGAGGATCTTACGGGCTGTGTGGCGGCGGAGAGAGAGAGAG 2038
 QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
 DB 2039 CGGCTCAGATGTTCTTGTGTAATACCAACAGACAGAGATTTGACCTCCAAAGGTTTAC 2098

QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 DB 2099 CCACCTGTCCCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2158
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
 DB 2159 GCTACAGAGCTCTCGTGTGAAGAGACTTGGGCACTTCAAGGTGAGACCTACATTCGCTCG 2218
 QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 DB 2219 GAGGTGTGTAGAGACTTCATACAGGTGGGATGTGGGACCAAGGGAGATATGTAC 2278
 QY 461 ValGlyAspValArgSerValIleThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
 DB 2279 GTGCGGAGCTCCGACGCTCACAGAGACATATCCAGAGTGGGCGCATTCGATCTG 2338
 QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2339 GTGATTTGGGGAGTCCCTGCAATGACCTCTCCATCGTCAACCTGCTCCGAGGGCCTC 2398
 QY 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
 DB 2399 TACGAGGACCTGGCGCGCTCTTGTGAGTTCTACCGCTCTGCAATGATGCGGCGCC 2458
 QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
 DB 2459 AAGGAGGAGATGATCCCTCTTCTTCTGCTCTTGAAGATGTGTGTGCGCATGGCGCTT 2518
 QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
 DB 2519 AGTGACAAAGGAGACATCTCGCATTTCTCGAGTCCAAACCTGTGATGATGAGCCAAA 2578
 QY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyLysLeuProGlyMetAspArg 580
 DB 2579 GAAGTGCACCTCACACAGGGCCCGCTACTTCTGGGTAACTTCCCGGTATGAACAGG 2638
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGlnGluCysLeuGlnHisGlyArg 600
 DB 2639 CCGTTGGCATCCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2698
 QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
 DB 2699 ATAGCCAAATTACAGAAAGTACGAGACCATTAACAGGTCAAACTCCATTAACAGGCG 2758
 QY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
 DB 2759 AAAGACCAAGATTTCTGTCTTCATGAATGAGAAAGAGAGACATTTATGTGTGCACTGAA 2818
 QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 DB 2819 ATGGAAAGGATATTTGGTTTCCAGTCCATATATGACCTCTCCACATGAGCCGCTTG 2878
 QY 661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
 DB 2879 GCGAGGAGAGAGACTGCTGGGCGGTCATGAGAGCGTGCACATGATCCGCCACTTCTGCT 2938
 QY 681 ProLeuLysGluTyrPheAlaCysVal 698
 DB 2939 CCGCTGAAGAGATATTTGTGTGTGTG 2995
 RESULT 8
 AF067972 3005 bp mRNA linear PRI 12-FEB-2001
 LOCUS Homo sapiens DNA cytosine methyltransferase 3 alpha (DNMT3A) mRNA,
 DEFINITION complete cds.
 ACCESSION AF067972
 VERSION AF067972.2 GI:12746531
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

QY 421 AAThrcGlyLeuLeuValLeuLysAspLeuGlyIleGluValAspArgTyrIleAlaSer 440
 DB 2159 GCTACAGGGCTCTGCTGCTGTAAGAGACTTGGGCACTCAGGTGACCGGCTACATTGCTCTG 2218
 QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 DB 2219 GAGGTGTGTGAGACTCCATCAAGGTGGCATGTGTGGCAGCAGGGAAGATCATGTAC 2278
 QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrGlyProPheAspLeu 480
 DB 2279 GTGCGGAGAGCTCCGACGCTCAACAGAAACATATCCAGAGGGGCCCATTCGATCTG 2338
 QY 481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2339 GTGATTGGGGGAGCTCCCTGCAATGACTCTTCATGTGTAAACCTGCTGCGAAGGGCTTC 2398
 QY 501 TyrGlnGlyThrGlyArgLeuPheGluPheGlyTyrArgLeuLeuHisAspAlaArgPro 520
 DB 2399 TACGAGGCACTGGCCGGCTCTTCTTGATTCTACCGCTCCTGATGATGGGGCCC 2458
 QY 521 LysGlnGlyAspAspArgProPhePheTyrPheGluAsnValValAlaMetGlyVal 540
 DB 2459 AAGGAGGAGATGATCGCCCTTCTCTGCTTGTGAGAAATGTGTGGCCATGGGCGTT 2518
 QY 541 SerAspLysArgAspIleSerArgPheLeuGlnSerAsnProValMetIleAspAlaLys 560
 DB 2519 AGTACAAAGAGGACATCTCCGATTTCTCGAGTCCAAACCCTGTGATGTATGATCCAAA 2578
 QY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyLysAsnArg 580
 DB 2579 GAAGTGTAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTCCGGTATGAACAG 2638
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGlnLeuGlnGlyCysLeuGlnHisGlyArg 600
 DB 2639 CCGTGGCATCCACTGATGATGATAGCTGAGCTGAGAGTGTCTGGAGCATGACAG 2698
 QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
 DB 2699 ATAGCCAAAGTTCAGAAAGTGAAGAGCACTTACTACAGAGTCAAACTCCATAAGAGAGG 2758
 QY 621 LysAspGlnHisPheProValPheMetAsnGlnLysGluAspIleLeuTyrCysThrGlu 640
 DB 2759 AAAGCCAGCATTTTCTCTGCTTCAATGATGAAGAAGAGCATTTATGGTGCATGA 2818
 QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 DB 2819 ATGAAAGGTAATTTGTTCCCACTCACTATCTGACGTCTCCAAACATGAGCGCTTG 2878
 QY 661 AlaArgGlnArgLeuGlnArgSerTyrSerValProValIleArgHisLeuPheAla 680
 DB 2879 GCGAGGCGAAGACTGCTGGCGGATCATGAGCGTGCAGCATCTCCGACCTCTTCCCT 2938
 QY 681 ProLeuLysGlnTyrPheAlaCysVal 689
 DB 2939 CCGCTGAAGAGTATTTTGGCTGTGTG 2965
 RESULT 9
 CS050242 3604 bp DNA linear PAT 23-MAR-2005
 LOCUS Sequence 26 from Patent WO2005021757.
 DEFINITION CS050242
 ACCESSION CS050242
 VERSION CS050242.1 GI:61889497
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 REFERENCE
 1 Tomme, P. H. and van Rompaey, L.
 Peptides and polynucleotides for use as a medicament
 Patent: WO 2005021757-A 26 10-MAR-2005;
 JOURNAL

FEATURES Galapagos Genomics N.V. (BE)
 source Location/Qualifiers
 1..3604
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3604
 Score: 3714.00 Matches: 679
 Percent Similarity: 98.8% Conservative: 2
 Best Local Similarity: 98.5% Mismatches: 8
 Query Match: 98.8% Indels: 0
 DB: 2 Gaps: 0
 US-10-623-813-85 (1-689) x CS050242 (1-3604)
 QY 1 MetAsnAlaValGluGlnAsnGlnAlaSerGlyGlnSerGlnLysValGluGluAlaSer 20
 DB 217 ATGAATGCTGTGGAGAAACCAAGGGCCCGGAGTCTCAGAAAGTGAAGAGGCGCAC 276
 QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 DB 277 CCTCTGCTGTGACAGCACCCCACTGACCCCGCATCCCACTGTGTGTACCAACGCTTAG 336
 QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 DB 337 CCGGTGGGGTCCGATGCTGGGGAGACAAAGATGCCCAAGAGCGATGACGAGCCAGAG 396
 QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGlnLeuValTyrGlyLysLeuArgGlyPhe 80
 DB 397 TACAGAGAGGCGCGGGCTTTGGCATTTGGAGCTGTGTGGGAAACTGGGGGCTTC 456
 QY 81 SerTyrTyrProGlyArgIleValSerTyrTyrMetThrGlyArgSerArgAlaAlaGlu 100
 DB 457 TCTGTGTGGCAGGCGCGCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 516
 QY 101 GlyThrArgTyrValMetTyrPheGlyAspGlyLysPheSerValValCysValGluLys 120
 DB 517 GGCACCGCGCTGGTCACTGT 576
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 DB 577 CTGATCCGCTGACCTGTTGT 636
 QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 DB 637 ATGTACCGCAAGGCATCTACAGAGTCTGTGAGGTGGCCAGCCCGCGGGGAAGCTG 696
 QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsn 180
 DB 697 TTCCCGGTGTCCACGACGAGATGAGATGACCTCCCAAGGCGGTGTGTGTGTGTGTGT 756
 QY 181 LysGlnMetIleGluTyrPheAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
 DB 757 AAGCCCAATGATGATGAGGCTTGGGGGCTTCCAGGCTTCTGTGCTTCAAGGGCTTGAAG 816
 QY 201 ProProGluGluGlnLysAsnProTyrLysGlnValTyrThrAspMetTyrValGluPro 220
 DB 817 CCACCAAGAGAGAGAAATCCCTTCAAGAAATGATACAGGACATGTGTGTGTGTGTGTGT 876
 QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
 DB 877 GAGGCACTGCTCTTACGACCACTCCACGCAAAAGCCCGAAGAGAGCAGCGGAG 936
 QY 241 LysProLysValLysGlnIleLysAspGluArgThrArgGluArgGluValTyrGluVal 260
 DB 937 AAGCCCAAGTCAAGAGATTTATGATGAGGCAAGAGAGCGGTGTGTGTGTGTGTGTGT 996
 QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 997 CCGGAGAAAGTCCGGAACATTTGAGGACATCTGATCTCTGTGTGTGTGTGTGTGTGTGT 1056

```

QY      281  LeuGluNH1SerProLeuPheH1leGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
DB      1057  CTGGAAACACCCCTCTTGGTGGAGGAATGTGCAGAAATCGCAAGAACTTTCTTGGAG 1116
QY      301  CysAlaIYrGlnIYrAspAspAspGlyYrGlnSerTYrCysThrIleCysCysGlyGly 320
DB      1117  TGTGGGTACAGATGACACAGCAGCGGCTACCGACTCTACTGCAACATCTGTGTGGGGAC 1176
QY      321  ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
DB      1177  CGTGAGGTCTCATGTGTGGGAAACAACTGTGCAAGGTGCTTTGGCTGGAGTGTGTG 1236
QY      341  AspLeuLeuValGlyProGlyAlaAlaAlaIleLysGluAspProTrpAsnCys 360
DB      1237  GACCTCTTGTGGGGCGGGGGGCTGCGCAGGACGACATTAGGAAGAACCTCTGGAATCTC 1296
QY      361  TYrMetCysGlyYrIleLysGlyYrGlyYrGlyLeuLeuArgArgGluAspTrpProser 380
DB      1297  TACATGTGGGGGACAAAGGGTACCTACGGGCTGTGGGGGCGGAGAGGACTGGCCCTCC 1356
QY      381  ArgLeuGlnMetPhePheAlaAsnAsnH1sAspGlnGluPheAspProTrpLysValTYr 400
DB      1357  CGGCTCCAGATGTTCTTCCCTAATTAACAAGACAGAAATTTGACCTCCAAAGCTTAC 1416
QY      401  ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
DB      1417  CCACCTGTCCAGCTGAGAAAGAGAAAGCCATCCGGGTCTGTCTCTTTGATGGAATC 1476
QY      421  AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTYrIleAlaSer 440
DB      1477  GGTACAGGGCTCTGTGTGCTGAAGACTTGGGGATTGAGGTGACCGCTACATTGCTCTG 1536
QY      441  GluValCysGluAspSerIleThrValGlyMetValArgH1sGlnGlyLysIleMetTYr 460
DB      1537  GAGGTGTGTGAGAGACTCCATACAGGTGGGATGTGTGGGACACGAGGAAGATCATGTAC 1596
QY      461  ValIleGlyAspValArgSerValThrGlnLysH1sIleGlnGlyIleProPheAspLeu 480
DB      1597  GTGGGGGACGTCCGCGAGCTCACACAGAAAGCAATATCAGAGAGGGGCCCATTCGATCG 1656
QY      481  ValIleGlyLysSerProCysAsnAspLeuSerIleValAspProAlaArgLysGlyLeu 500
DB      1657  GTGATTTGGGGGAGCTGCTGCAATGACCTTCATGATGTAACCTCTGCGAAGGGCTTC 1716
QY      501  TYrGluGlyThrGlyArgLeuPheGluPheTYrArgLeuLeuH1sAspAlaArgPro 520
DB      1717  TACGAGGGGACTGGCCGCTCTTTGAGTTCTACCGCTCTGATGATGCGCGGCC 1776
QY      521  LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaAlaMetGlyVal 540
DB      1777  AAGGAGGAGATGATGCGCCCTCTTCTGTGAGATGTGGTGGGCTGAGGCGGCTT 1836
QY      541  SerAspLysArgAspIleSerArgPheLeuGlnLysSerAspProAlaMetIleAspAlaLys 560
DB      1837  AGTGAAAGAGGACATCTCGCATTTCTGAGCTCAACCTCTGATGATGATGCGCAA 1896
QY      561  GluValSerAlaAlaH1sArgAlaArgTYrPheTrpGlyAsnLeuProGlyMetCysArg 580
DB      1897  GAAGTCTACGCTGCAACAGGGCCCGCTACTTGTGGGTAACTTCCCGTTAGAACAG 1956
QY      581  ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluH1sGlyArg 600
DB      1957  CCGTTGGCATCTCACTGTGAATGATAGCTGAGACTGAGAGAGTGTCTGGAGCATGGAGG 2016
QY      601  IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
DB      2017  ATAGCCAAAGTTTCAGAAAGTGAAGCACTTACTACAGGTCAAACTTCATTAAGCAGG 2076
QY      621  LysAspGlnH1sPheSerValPheMetAsnGluLysGluAspIleLeuTrpCysTrpGlu 640
DB      2077  AAAGACACAGATTTTCTGTCTTCATGATGAGAAAGAGACATCTTATGTGTCACTGAA 2136

```

```

QY      641  MetGluArgValIlePheGlyPheProValH1sTYrThrAspValSerAsnMetSerArgLeu 660
DB      2137  ATGGAAGAGGATTTGTTTCCCACTGCACTTACTGACGTCTCAACATAGAGCGGCTTG 2196
QY      661  AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgH1sLeuPheAla 680
DB      2197  GCGAGCAGAGACTCTGCTGGCCGGTCATGAGAGCGTGCACATCATCGCACCTCTTCGCT 2256
QY      681  ProLeuLysGlyTYrPheAlaCysVal 689
DB      2257  CCGTGAAGAGATTTTGGCGTGTGTG 2283

RESULT 10
AF331856      4258 bp      mRNA      linear      PRI 02-JAN-2002
LOCUS        Homo sapiens DNA cytosine methyltransferase 3 alpha (DNMT3A) mRNA,
DEFINITION   complete cds.
ACCESSION    AF331856
VERSION      AF331856
KEYWORDS     SOURCE
ORGANISM     Homo sapiens (human)
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Homidae; Homo.
REFERENCE    1 (bases 1 to 4258)
             Ni,J., Pradhan,S. and Roberts,R.J.
             Cloning, expression and characterization of human DNMT3 genes
             UNPUBLISHED
             2 (bases 1 to 4258)
             Ni,J., Pradhan,S. and Roberts,R.J.
             Direct Submision
             Submitted (22-DEC-2000) New England Biolabs, 32 Tozer Road,
             Beverly, MA 01915, USA
FEATURES     source
             1..4258
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /chromosome="2"
             1..4258
             /gene="DNMT3A"
             /gene="DNMT3A"
             /codon_start=1
             /product="DNA cytosine methyltransferase 3 alpha"
             /protein_id="AA157039.1"
             /translation="MPSSGPGDTSSSAAREBERDRKDGEQEBPRKEERQEBSTARK
             VGRPGRRKHPPVESGDTPKDPAVISKSPMSMQDSGASLLPNGDLEKSEFQPEBGS
             PAGGCGKGAAPAEGBGAETLPEASRAVENSGCTPREGGAPEAGKEQKETNIESMKM
             ESSRRRLKGLUESLSLORPPRLTTPQAGDPYYSKRDRDWLARWKEARKKAKVI
             AGMANVAENOGGESQVVEASPPAVQOPTDPAFTVATTPPVSDADQKATACG
             DEPEVEDRGFGIGELVWGKLGFPFWBGRIVSMWMTGSRABEGTRVMWFGDGFFS
             VVVCVKLMPISFCSAFQATYKQPMYRKALYEVLOYASRAGKLFVCHOSDESDT
             AKAVIVONKPMIEMALGGFOPSGPKGLEPPEBKRPYKEVYDMMVPEPAAAYAPP
             AKRPKSTRAEKRYKVIIDTERRRLYVVRKKNIEDICISGSLNTLTLHPLEVG
             GNCQKCKFCLECAVOYDDGYSQYCTTCGGRBYLMGNNCCRCFCFCEVDLIVP
             GAAQAAIKEDPNWNCYMKGTGYGLRRREDPRLQNFANNHODEPRLPYRPPV
             AEKRXPRIYLSFDIATGLAVLKDLQVQDYIASVEEDSIITVGWRLQHLIMVG
             DVRSYTKHILQEMGFDFLVITGSPCNDSIYNPARKGLYEGKRLFEEFVRLGLHMG
             KEGDRPFWMLPENYVAMGVSDKRDISFLESNPVMIDAKEVSAHARAYFWGNLPGM
             NPLASTYNDKLELOECLEHGRIAKFSKVRITTSNSIKQCKDGHFVFNKEXEDIL
             KCTERERPFGPIVHTDYSNMRRLARORLGRSWSVPVIRHLFAPLKEYFACV"

ALIGNMENT SCORES:
Pred. No.: 0
Score: 3714.00
Percent Similarity: 98.84
Best Local Similarity: 98.84
Query Match: 5
DB: 5
Length: 4258
Matches: 679
Conservative: 2
Mismatch: 8
Indels: 0
Gaps: 0

```

US-10-623-813-85 (1-689) x AF331856 (1-4258)

OY 1 MetAsnAlaValAGluGluAsnGlnAlaSerGlyGluSerGlnLysValAGluGluAlaSer 20
 DB 887 ATGAATGCTGCTGGAGAAAAACAGGGGCCCGGGAGCTCAGAAAGGTGAGGAGGCGCAC 946
 OY 21 ProProAlaValAGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 DB 947 CTTCTGCTGCTGACAGCCCACTGACCCCGCATCCCACTGTGCTACACAGCCTGAG 1006
 OY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 DB 1007 CCCGTGGGGTCCGAGTCTGGGGACAAGAAATGCCCAAAAGCAGCGCATACAGCCAGAG 1066
 OY 61 TyrGluAspGlyArgGlyPheGlyTyrGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80
 DB 1067 TACGAGGACGGCCGGGGCTTTGGCATTTGGGAGCTGTGTGGGGGAAACTCGCGGGCTTC 1126
 OY 81 SerTPRTPProGlyArgGlyLeuValSerTPRTPMetThrGlyArgSerArgAlaAGlu 100
 DB 1127 TCTGTGTGCGCAGCGCGCATTTGTCTTGTGTGATGACGGCGCGAGCCGACAGCTGAA 1186
 OY 101 GlyThrArgTyrValMetTyrPheGlyAspGlyLysPheSerValValCysValGluLys 120
 DB 1187 GGCACCCCGCTGGGTATGTGTGTCGAGAGCGGCAAAATCTCACTGTGTGTGTGAGAA 1246
 OY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 DB 1247 CTGATGGCGGTAGCTCGTTTGTGACGTGGCTTCCACAGGCGCACGTAACAAGCAGCCC 1306
 OY 141 MetTyrArgLysAlaTyrGlyGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 DB 1307 ATGTAACCCCAAGGCATCTACAGAGTCTTCGAGGTGGCCAGACGCCCGCGGAGAGCTG 1366
 OY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValAGluValGlnAsn 180
 DB 1367 TTCCCGGTGTGCGACAGCACGCAATGAGTGAACCTGCCAAGCCCTGTGAGGTGCAAG 1426
 OY 181 LysGlnMetTyrGlnTyrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGln 200
 DB 1427 AAGCCCATGATGATGAGGCGCTCGGGGGGCTTCCAGCCTTCTGGCCCTAAGGCGCTG 1486
 OY 201 ProProGluGluGlnLysAsnProTyrLysGluValTyrThrAspMetTyrValGluPro 220
 DB 1487 CCACAGAGAGAGAGAGAAATCCCTACAAAGAGTACACGAGACATGTGGGTGAACT 1546
 OY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
 DB 1547 GAGGAGCTGCTTACGACCACTTCCACAGCCCAAAAGCCCGGAGAGGACACAGCGAG 1606
 OY 241 LysProLysValLysGluLysLeuAspGluArgThrArgGluArgLeuValTyrGluVal 260
 DB 1607 AAGCCCAAGGTCAAGAGATTAATGATGAGCCACAAAGAGACGGCTGTGTAACGAGTG 1666
 OY 261 ArgGlnLysCysArgAsnTyrGluAspTyrLysSerCysGlySerLeuAsnValThr 280
 DB 1667 CGGCGAAGTGCAGAACTTGAAGACATCTGATCTCTGTGGGAGCCTCATGTATAC 1726
 OY 281 LeuGlnHisProLeuPheHisGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
 DB 1727 CTGGAACACCCCTCTTCTGTGAGAGAAATGTCCCAAAATGCAAGAACTCTTCTGAG 1786
 OY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrTyrLysCysGlyGly 320
 DB 1787 TGTGGGTACAGTACGACGACGCGCTACAGTCTTACTGACCATCTGTGTGGGGAGC 1846
 OY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValAGluCysVal 340
 DB 1847 CGTGGGTGCTCATGTGCGGAACAACATCTGTGAGGTGCTTTTGTGCTGAGAGTGTG 1906
 OY 341 AspLeuLeuValAGlyProGlyAlaAlaGlnAlaAlaTyrGlyLysProTyrAsnCys 360

DB 1907 GACCTTGTGTGGGGCCGGGGGCTGCCAGGACGACCATTAAGAAAGACCCCTGAACTGC 1966
 OY 361 TyrMetCysGlyHisLysGlyTyrThrTyrGlyLeuLeuArgArgGluAspThrProSer 380
 DB 1967 TACATGTGCGGGCACAAGGGTACTTACGAGCTGTGCGCGCGCGAGAGGATGCGCCCTC 2026
 OY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
 DB 2027 CGGCTCCAGATGTTCTTGTGCTTAATACACAGCAGAAATTTGACCTTCCAAAGTTTAC 2086
 OY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyLe 420
 DB 2087 CCACCTGTCCCGAGCTGAGAAAGAGAACCCATCCGGGTGCTGTCTCTTATGAGAAATC 2146
 OY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyTyrGlnValAspArgTyrTyrLeuSer 440
 DB 2147 GCTACAGGGCTCTGTGTCTGAGAGCATTTGGGATTCAGGTGACCGGCTACATTCCTG 2206
 OY 441 GluValCysGluAspSerTyrThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 DB 2207 GAGTGTGTGAGACTCCATCACGCTGGCATGTGTGCGGACCAAGGGAAGATCATGTAC 2266
 OY 461 ValGlyAspValArgSerValThrGlnLysHisGlnGlnTyrGlyProPheAspLeu 480
 DB 2267 GTGGGAGAGTCCGACGCTCACACAGAGCATATCCAGAGGTGGGCCCATTCGATCTG 2326
 OY 481 ValTyrGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2327 GTGATTTGGGGAGAGTCCCTGCAATGACTTCCATGTCTCAACCTGTGTGCAAGGGCTC 2386
 OY 501 TyrGlnGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
 DB 2387 TACGAGGCACTGGCGGCTCTTCTTGAATTCACCCCTCTCTCATATGTCGGGCCC 2446
 OY 521 LysGlnGlyAspAspArgProPhePheThrLeuPheGluAsnValValAlaMetGlyVal 540
 DB 2447 AAGAGGAGATGATGCGCCCTTCTTGTGCTCTTGTGAATGTGTGGCCCATTCGGCGCTT 2506
 OY 541 SerAspLysValArgAspTyrLysThrPheLeuGluSerAsnProValMetLysAspAlaLys 560
 DB 2507 AGTGCAAGAGGAGATCTCCGATTTCTCGAGTCCAAACCTGTATGATGATATGCAAA 2566
 OY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArg 580
 DB 2567 GAGGTGACGTGCAACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGTATGAACAG 2626
 OY 581 ProLeuAlaSerThrValAsnAspLysLeuGlnGlnCysLeuGlnHisGlyArg 600
 DB 2627 CCGTTGGCATCCACTGTAATGATTAAGCTGAGAGCTGCAGAGGTGTCTGGACATGGCAGG 2686
 OY 601 IleAlaLysPheSerLysValArgThrTyrThrThrArgSerAsnSerTyrLysGlnGly 620
 DB 2687 ATAGCAAGTTCACGAAAGTGAAGCACTTACTAGAGTCAAACTCATTAAGACAGGGGC 2746
 OY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspTyrLeuTyrCysThrGlu 640
 DB 2747 AAGACCGCATTTTCTCTGTCTTCAATGATGAGAAAGAGCATTTATGTGTCATGAA 2806
 OY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 DB 2807 ATGAAAGGATATTGTTGTTCCAGTCCACTATATGACGTCTCCAAATGAGCGGCTTG 2866
 OY 661 AlaArgGlnArgLeuLeuGlyArgSerThrProValIleLeuArgHisLeuPheAla 680
 DB 2867 GCGAGGCAAGAACTGTGGCGGTGATGAGCGGTGCCAGTATCCGCCACTTTCGCT 2926
 OY 681 ProLeuLysGluTyrPheAlaCysVal 689
 DB 2927 CCGCTGAAGAGATTTTGGCGGTGTG 2953

RESULT 11
 DD171352
 LOCUS DD171352 4294 bp DNA linear PAT 19-DEC-2005

DEFINITION Marker for diagnosing human subarachnoid hemorrhage and its use.
 ACCESSION DDI171352
 VERSION DDI171352.1 GI:83959786
 KEYWORDS JP 2005151854-A/10.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 4294)
 AUTHORS Yamaura, A., Kubota, M., Machida, T., Uchino, F., Kobayashi, E., Saeki, N., Takiguchi, M. and Hiwasa, T.
 TITLE Marker for diagnosing human subarachnoid hemorrhage and its use
 JOURNAL Patent: JP 2005151854-A 10 16-JUN-2005;
 Japan Science and Technology Agency
 COMMENT OS Homo sapiens
 PN JP 2005151854-A/10
 PD 16-JUN-2005
 PF 21-NOV-2003 JP 2003393161
 PI akira yamaura,mocoo kubota,roshio machida,fukuo uchino,eichi
 PI kobayashi,
 PI nackatsu saeki,masaki takiguchi,takaki hiwasa CC
 FH Key Location/Qualifiers
 FEATURES
 source 1..4294
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 4294
 Score: 3714.00 Matches: 679
 Percent Similarity: 98.8% Conservative: 2
 Best Local Similarity: 98.5% Mismatches: 8
 Query Match: 98.8% Indels: 0
 Gaps: 0
 DB: 2
 US-10-623-813-85 (1-689) x DDI171352 (1-4294)
 QY 1 MetAsnAlValGluGluAsnGlnAlaSerGlyGluSerGlnValGluGlnAlaSer 20
 DB 907 ATGAATCGCTGTGCAAGAAACCAAGGGCCCGGGAGCTCAGAAAGTGGAGCGCCAGC 966
 QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 DB 967 CCGCTCGTGTGACGACGCCACTGACCCCGCATCCCGCATCTGTGTACACGCCCTGAG 1026
 QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 DB 1027 CCGGTGGGTCCGATGCTGGGACAGAAATGCCACCAAGCAGCGATGACGAGCCAGG 1086
 QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrrpGlyLysLeuArgGlyPhe 80
 DB 1087 TAGAGAGACGGCGGGGCTTTGGCATTGGGAGCTGTGTGGGGAACCTGGGGGCTTC 1146
 QY 81 SerTrrpProGlyArgGlyIleValSerTrrpTrrpMetThrGlyArgSerArgAlaAlaGlu 100
 DB 1147 TCCTGTGGCCGACCGCATTTGTCTTGTGTGATGACGGCGGAGCCGACGACGCTGAA 1206
 QY 101 GlyThrArgTrrpAlaMetTrrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
 DB 1207 GGCACCCGCGGGATCATGTGTTGGAGAGAGCAAAATTCATGATGTGTGTGAGAG 1266
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 DB 1267 CTGATCCCGCTGAGCTCGTTTTCAGATGGCTTCACAGGCGCACTGACAAAGCAGCCC 1326
 QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 DB 1327 ATGTACCGCAAAACCATCTACGAGGCTCTCAGGTGGCCAGCGCCCGGGGAAAGCTG 1386
 QY 161 PheProAlaCysHisAspSerAspGluSerGlyLysAlaValGluValGlnAsn 180

DB 1387 TTCCCGTGTCCACGACGCGATGAGTGAACCTGCCAAGCCGCTGAGGTGCAGAAC 1446
 QY 181 LysGlnMetIleGluTrrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
 DB 1447 AAGCCCATGATTAATGAGCCCTCGGGGGGCTTCAGCCCTTCTGGCCCTTAAGGCCCTGGAG 1506
 QY 201 ProProGluGluGluLysAsnProTrrpLysGluValTyrThrAspMetTrrpValGluPro 220
 DB 1507 CCACCAAGAAAGAGAAAGATCCCTACAAAGAGTACACGACATGTGGTGGAACT 1566
 QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
 DB 1567 GAGGAGCGTGGCTTACGACACCTCCACGCAAAAGACCCCGGAAGACACAGCGGAG 1626
 QY 241 LysProLysValLysGluIleLeuAspGluArgThrArgGluArgLeuValTyrGluVal 260
 DB 1627 AAGCCCAAGTCAAGAGATTATGTGAGCCACAAAGAGACCGCTGTGTGACGAGTGT 1686
 QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 1687 CGCAGAAAGTCCGGGAACATTGAGACATCTGCATCTCTGTGGAGCTCAATGTACC 1746
 QY 281 LeuGluHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
 DB 1747 CTGAAACACCCCTCTCTGCTGGAGGAATGCGCAAAATCGCAAGAACTGCTTCTGGAG 1806
 QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
 DB 1807 TGTGCGTACAGTACGACGACGACGCGCTACAGTCTTACTGACCACTGTGTGTGGGGC 1866
 QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
 DB 1867 CGTAGAGTCTAATGTGCGGAACAAACAATCTGTCAGGTGCTTTGGCTGAGTGTGTG 1926
 QY 341 AspLeuLeuValGlyProGlyValAlaGlnAlaAlaIleLysGluAspProTrrpAsnCys 360
 DB 1927 GACCTCTGTGGGGCGGGGCTGCGCAGGACGACATTAAAGAAACCCCTGGAACCTGC 1986
 QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGluAspTrrpProSer 380
 DB 1987 TACATGTGGGGCACAAAGGATCTTAACGCGGTGCTGGCGGGCGAGAGAGATGGCCCTCC 2046
 QY 381 ArgLeuGlnMetPhePheHisAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
 DB 2047 CGGCTCAATGTTCTTCTGCTTAATTAACGACGACGAAATTTGACCTCCAAAGTTTAC 2106
 QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 DB 2107 CCACCTGTCACAGCTGAGAAAGAAAGCCATCCGGGTGCTGTCTCTTGTGATGAATC 2166
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTrrpIleAlaSer 440
 DB 2167 GCTACAGGCTCTCTGTGTGAAGACTTGGGCACTTGAGTGAACGCTACATTTGCTTCG 2226
 QY 441 GluValCysGlyAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 DB 2227 GAGGTGTGAGGACTTCATCAACGTTGGCATGTGTCGACCAAGGAAAGATCACTTAC 2286
 QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrrpGlyProPheAspLeu 480
 DB 2287 GTCCGGGAGTCCGCGAGCTCACAGAGACATATCAAGAGTGGGCCCATTCGATCTG 2346
 QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2347 GTGATTTGGGGGAGTCCCTGCAATGACCTCTCATGTGAACCTGTGCGCAAGGCGCTC 2406
 QY 501 TyrGluGlyThrGlyArgLeuPhePheTyrArgLeuLeuHisAspAlaArgPro 520
 DB 2407 TACAGGAGCATGGCGGCTTCTTTGAGTTCTTACGCTCTCTGATGATGCGGGCCC 2466
 QY 521 LysGluGlyAspAspArgProPhePheTrrpLeuPheGluAsnValAlaMetGlyVal 540

Db 2467 AAGGAGGAGATGATGCGCCCTTCTTCTGGCTTTGAGATGTCGTCGCGCATGCGGCTT 2526
 QY 541 SerAPlyeaArgAsp111SerArgPheLeuGluserAsnProValMet111eaSPAlaLys 560
 Db 2527 AGTGACAGAGGAGACATCTCGCATTTCTCGAGTCCAAACCTGTGATGATGCCAA 2586
 QY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
 Db 2587 GAAGTGTACAGTCCACACAGGGCCGCTACTCTTGGGGTAACTTCCCGGTATGAACAGG 2646
 QY 581 ProLeuAlaSerThrValaAsnAspLysLeuLeuGlnGluCysLeuGlnIun1sglyArg 600
 Db 2647 CGTTGGCATCCACTGTGAATGATTAAGCTGGAGCTGAGAGAGTGTCTGAGCATGCGCAGG 2706
 QY 601 11eAlaLysPheSerLysValaArgThr11eThrThrArgSerAsnSer11elyGlnGly 620
 Db 2707 ATAGCCAAAGTTCAGCAAGAGGAGCATTACTACGAGTCAAACTCCATMAAGCAGGGC 2766
 QY 621 LysAspGlnHisPheProValaPheMetAsnGluLysGlnAsp11eLeuTyrPySThrGlu 640
 Db 2767 AAGAGCAGCATTTTCTCTGTCTTCATGATGAGAAAGAGACATCTTATGTGTCACTGAA 2826
 QY 641 MetGluArgValPheGlyPheProValaHisTyrThrAspValSerAsnMetSerArgLeu 660
 Db 2827 ATGGAAAGGATATTGGTTTCCAGTCCACTATAGTACGCTCCCAACATGAGCCGCTTG 2886
 QY 661 AlaArgGlnaArgLeuGlyArgSerTyrSerValProVal11eArgHisLeuPheAla 680
 Db 2887 GCGAGCAGACAGATCGTCGGCCCGCTCATGAGCGTGCAGCTCAGCCACCTTCCTGCT 2946
 QY 681 ProLeuLysGluTyrPheAlaCysVal 689
 Db 2947 CCGCTGAGAGATATTGCGTGTGTG 2973

RESULT 12 4294 bp mRNA linear PRI 30-JUN-2004
 LOCUS BC043617
 DEFINITION Homo sapiens DNA (cytosine-5'-methyltransferase 3 alpha,
 transcript variant 3, mRNA (cDNA clone MGC:50948 IMAGE:6150112),
 complete cds.
 ACCESSION BC043617 GI:27694443
 VERSION BC043617.1
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE 1 (bases 1 to 4294)
 Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 Diachenko L., Martusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stachler M., Soares M.B., Donald T.F., Casavant T.L.,
 Scheer T.E., Brownstein M.J., Usdin T.B., Teshiyuki S.,
 Carinci P., Prange C., Kana S.S., Loquellano N.A., Peters G.J.,
 Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J.,
 McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
 Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S.,
 Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,
 Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
 Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butlerfield Y.S., Krzywinski M.I., Skalska U., Smalins D.E.,
 Scherch A., Schein J.E., Jones S.J. and Marra M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 PUMED 12477932
 REFERENCE 2 (bases 1 to 4294)
 AUTHORS Strausberg R.

TITLE JOURNAL
 REMARK COMMENT
 Direct Submission
 Submitted (09-JAN-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Offices, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC/DCID/DTF
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcsc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Buterfield,
 Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
 Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao,
 Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
 Telka Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prabhu,
 Paraneh Saeedi, JR Santos, Angelique Scherch, Ursula Skalska,
 Duane Smalins, Jeff Stott, Miranda Teai, George Yang, Jacque
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAX Plate: 89 Row: 0 Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.
 FEATURES
 source
 Location/Qualifiers
 1..4294
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:50948 IMAGE:6150112"
 /tissue_type="Skin, melanocytic melanoma."
 /clone_lib="NIH MGC_72"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 1..4294
 /gene="DNMT3A"
 /note="synonyms: DNMT3A2, M.Hsa111A"
 /db_xref="GeneID:1788"
 /db_xref="MIM:602769"
 238..2976
 /gene="DNMT3A"
 /codon_start=1
 /product="DNA cytosine methyltransferase 3 alpha, isoform
 a"
 /protein_id="AAH43617.1"
 /db_xref="GI:27694444"
 /db_xref="GeneID:1788"
 /db_xref="MIM:602769"
 /translation="MPAMPSSGCPDSTSSAAREEDRKDGEDEBGRKEERQEPSTT
 AKVGRGPKRGKRPVSGDTPKPAVSKSSMODSSASELLPAGDEKSEPOPE
 EESPAGGKGAPABEGEAETLPKASRAVEGCCPKRGAPAGAEKQETNIES
 NMGESRGRLRGGLMESLSTLRPMPRLTPOAGDPVYSKRKDEMLARKEAEKA
 KTIAGNNAAVEENGQGESEOKVEASPPAVQPTDPASPVVATTPPEVSGDADKAKT
 AADDPENNAEDGFGFGLVWKLKLFVSWPRLIVSWMTGSRRAEGTRVWVWFEDG
 KRSVAVCEVLEMPLEISFCSAFHOATYKQPKYKALYEVLQVASSRAGKLPVCHDSE
 SSTAAVAEYONKEMLEALGFGPQSPGKLEPSEKPNYKXEVYTDWMEPEAAVAP
 PPAPKKRSTKREKVEKILDERTRERLVYEVROKRNKYEIVTDMWEPAAVAP
 PPGKCNKQKQKTECAVOYDDGYOSYCTTCGGREVMCCNNCCRGCEVCVUHL
 VPGQAQAALIKEDPNCYCKGKTYGLILAREDEPMSYOMEPANNDPEPDKYIP
 EPPAPKRRKRIPLSLFDGATLVLKVDIGIVDRYIASVEVDSITVGMHQKQIM
 VYGDVRSYQKIQEMGEPDVLIGSPKNDLSIVNPARGLYEGTGLTFEYRLIHD
 AAPKGDPRPFPMLENNVAMGVSDKRDLSRPLESPNPMIDAKYVSAHARARFVWNL
 PPMNRPLASTVNDKLEQCELEHGRILAKSKRTITRTRENSIKQGDQFVPMENKE
 DILKCTEMERVGPVHTVDVSNMRRLARQRLGSSWSVPVIRHLPAFLKEYFACV"

ORIGIN
 Alignment Scores:
 Pred. No.: 0
 Length: 4294

Score:	3714.00	Matches:	679
Percent Similarity:	98.8%	Conservative:	2
Best Local Similarity:	98.5%	Mismatches:	8
Query Match:	98.8%	Indels:	0
DB:	5	Gaps:	0
US-10-623-813-85 (1-689) x BC043617 (1-4294)			
QY	1	MecAsnAlaValGluGluAsnGlnAlaSerGlyGluSerGlnValGluGlnAlaSer	20
DB	907	ATGAAATCTGTGGAAAGAAAACAGGGCCCGGGAGTCTCAGAAAGTGGAGAGCCAGC	966
QY	21	ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu	40
DB	967	CTCTCTGCTGGACGACGCCACTGACCCGACATCCCACTGGCTGCTACCAAGCCTGAG	1026
QY	41	ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu	60
DB	1027	CCGGTGGGTCGCAATGCTGGGGAACAGATGCCAAGAGAGCGCATGACGAGCCAGAG	1086
QY	61	TyrGluAspGlyArgGlyPheGlyIleGlyGluValIleThrGlyLysLeuArgGlyPhe	80
DB	1087	TACGAGACCGCGCGCTTGGCATTGGGAGCTGGTGTGGGGAACCTGCGGCGCTTC	1146
QY	81	SerTrpTrpProGlyArgGlyIleValSerTrpTrpMetThrGlyArgSerArgAlaGlu	100
DB	1147	TCTCGTGGCCAGGCCCCCATTTGTCTTGGTGAAGAGGCGGAGCCAGACGACTGAA	1206
QY	101	GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys	120
DB	1207	GGCACCCGCTGGGTCAATGCTGTTGGAGAGCGGCAAAATTCACAGTGTGTGTGAGAG	1266
QY	121	LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro	140
DB	1267	CTGATGCGCTGAGCTCTGTTTTCAGTGTCTTCCACAGGCGCAAGTCAACAGCACCCC	1326
QY	141	MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu	160
DB	1327	ATGTACCGCAAGACCATCTACAGAGTCTCTGAGGTGCGCACGCCGCGGGGAGCTG	1386
QY	161	PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsn	180
DB	1387	TTCGCGGTGTGCACGACGACGATGAGTGAACACTGCCAAGGCGGTGAGTGCAGAAC	1446
QY	181	LysGlnMetIleGluTyrPalaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu	200
DB	1447	AACCCCATGATTGATGGGCCCTGGGGGCTTCCAGCTTCTGGCCCTTAAGGGGCTGGAG	1506
QY	201	ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro	220
DB	1507	CCACCAAGAAAGAAAGAAATCCCTACAAAGAAAGTGAACGACATGTGGTGGAACTT	1566
QY	221	GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu	240
DB	1567	GAGGACGCTGCTACGACCACTCCACAGCCAAAGCCCGGAGAGACACAGCGGAG	1626
QY	241	LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValIleTyrGluVal	260
DB	1627	AACCCCAAGGTCAAGGATTAATTGATGACGACCAAGAGAGCGGCTGGTGTACAGGTG	1686
QY	261	ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr	280
DB	1687	CGGCAAGAGGCGGAACTTGAAGACATCTCATCTCTGTGGAGCTTCATATTTACC	1746
QY	281	LeuGlnHisProLeuPheIleGlyLysMetCysGlnAsnLysCysLysAsnCysPheLeuGlu	300
DB	1747	CTGGAACACCCCTCTTGTGGAGGAAATGTGCCAAACTGCAAGAACTCTTCTGAG	1806
QY	301	CysAlaIleTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly	320
DB	1807	TGTGCGTACAGTACGACGACGAGGCTTACCATGCTTCAATCTCTGTGGGGGCT	1866
QY	321	ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal	340

DB	1867	CTGAGGTCTCATCTGCGGAAACAAACACTGTCTGAGGTGCTTTGCGTGGAGTGTG	1926
QY	341	AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProThrAsnCys	360
DB	1927	GACCTCTGTGGGGCCGGGGGCTGCCAGGACCATTAAGGAACCCCTGGAACCTG	1986
QY	361	TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGluAspTrpProSer	380
DB	1987	TACATGTGGGCAAGGATACCTACGAGCTGTGCGGCGGCGGAGAGAGCTGGCCCTCC	2046
QY	381	ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr	400
DB	2047	CGGCTCAGATGTTCTTCCCTTAATACACAGACAGGAATTTGACCTCCAAAGTTTAC	2106
QY	401	ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle	420
DB	2107	CCACTGTCCCACTGAGAAAGAGAACCCATCCGGGTGCTGTCTCTTGTGATGGAATC	2166
QY	421	AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer	440
DB	2167	GCTACAGGCTCTGCTGTGAAGACTTGGGCATTCAAGTGAACGCTACATTTGCTCG	2226
QY	441	GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr	460
DB	2227	GAGGTGTGAGAGACTCCATCACGCTGGCATGTGTCGACACAGGGAAGATCATGTAC	2286
QY	461	ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrGlyProPheAspLeu	480
DB	2287	GTCGGGAGCGTCCGCGCTGCACACAGAAACATATCCAGAGAGGGGCCCATTCGATCTG	2346
QY	481	ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu	500
DB	2347	GTAATTGGGGAGATCTCTGCAATGACTTCTCATGTCAACCTCTGCGCAAGGCTTC	2406
QY	501	TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro	520
DB	2407	TACAGAGGACATGCGCGCTCTTCTTGAATTCACCGCTCTGATGATGCGCGGCC	2466
QY	521	LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal	540
DB	2467	AAGGAGGAGGAGATGCGCCCTTCTTGCTTGTGAAGATGTGTGGCCCAAGGGCGTT	2526
QY	541	SerAspLysArgAspIleSerArgPheLeuGlnLysSerAsnProValMetIleAspAlaLys	560
DB	2527	AGTGAAGAGGACATCTCGGCAATTTCTCGAGTCCAACTGTGATGATGATGCCAA	2586
QY	561	GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg	580
DB	2587	GAAGTCTCAGTGCACACAGGGCCGCTACTTCTGGGGTAACTTCCCGGTATGAACAG	2646
QY	581	ProLeuAlaSerThrValAsnAspLysLeuGlnGluCysLeuGlnHisGlyArg	600
DB	2647	CCGTGGCATTCACCTGTGATATGATGAGCTGACAGAGTGTCTGAGAGCATGGCAAG	2706
QY	601	IleAlaLysPheSerLysValArgThrIleThrArgSerAsnSerIleLysGlnGly	620
DB	2707	ATAGCCAACTTCAGCAAGATGAGACCATTTACTACAGGTCAAATCTCAATTAAGCAGGG	2766
QY	621	LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu	640
DB	2767	AAAGACACACATTTCTCTCTTCAATGAATGAAGAAAGACATCTTAATGTGCACTGAA	2826
QY	641	MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu	660
DB	2827	ATGGAAGGATATTTGGTTTCCAGTCCACTATACGATGTCTCCAAACATGAGCCCTTG	2886
QY	661	AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla	680
DB	2887	GCGAGGCAAGACTGTGGGCGGCTCATGGAAGCGTGCCAGTATCCGCCACCTCTTGCT	2946
QY	681	ProLeuLysGluTyrPheAlaCysVal	689

Db 2947 CCGCTGAGAGATATTTGCGTGTG 2973

RESULT 13
CS050241 4395 bp DNA linear PAT 23-MAR-2005

LOCUS CS050241
DEFINITION Sequence 25 from Patent WO2005021757.
ACCESSION CS050241
VERSION CS050241
KEYWORDS GI:61889496

SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE
1 Tomme, P. H. and van Rompaey, L.
AUTHORS Polypeptides and polynucleotides for use as a medicament
TITLE Patent: WO 2005021757-A 25 10-MAR-2005;
JOURNAL Galapagos Genomics N.V. (BE)
FEATURES
source Location/Qualifiers
1..4395
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 4395
Score: 3714.00 Matches: 679
Percent Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatches: 8
Query Match: 98.8% Indels: 0
DB: 2 Gaps: 0

US-10-623-813-85 (1-689) x CS050241 (1-4395)

QY 1 MetAsnAlaValAGluGluAsnGlnAlaSerGlyGluSerGlnValAGluGluAlaSer 20
Db 1008 ATGAAATGCTGTGGAGAAACCAAGGGGCGGGGAGATCTCAAGAGTGAAGAGGGCCAGC 1067

QY 21 ProProAlaValAGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db 1068 CCTCGTGTGTCGAGAGCCAGCTGACCCCGCATCCCTGCTGCTACAGCTGAG 1127

QY 41 ProValAGlyValAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
Db 1128 CCGGTGGGCTCCGATGCTCTGGGGGACCAAGATGCCACCAAGCAGCGCATGACGCGCAGAG 1187

QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80
Db 1188 TAGAGAGAGCGCGGGGCTTTGGCATTTGGAGCTGTGTGGGAACTGCGGGGCTTC 1247

QY 81 SerTyrTrpProGlyArgIleValSerTyrTrpMetThrGlyArgSerArgAlaAlaGlu 100
Db 1248 TCCTGGTGGCCAGGGCGCATTTGTGTCTTGTGTGATACCGGGCGGAGCCGAGAGCTGAA 1307

QY 101 GlyThrArgTyrValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
Db 1308 GGCACCCCGCTGGGTCTGTGTGTCCGAGACGGCAAAATCTCAGTGTGTGTGTGAGAG 1367

QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db 1368 CTGATGCGCCGTGAGCTGTTTGGCATGTGGCTTCACAGGCGCATGACAAAGAGAGCC 1427

QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db 1428 ATGTACCGCAAAAGCCATACGAGGTCCTGCAAGTGGCAGCGCGCGGGAAGCTG 1487

QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAn 180
Db 1488 TTCCCGGTGTGCCACACAGCGATGAGATGACCGCAAGCCCTGTGAGGTCGCAAGAC 1547

QY 181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200

Db 1548 AAGCCCATGATGAATGGGCCCTGGGGGGCTTCAGACCTTCTGGCCCTAAGGCGCTGAG 1607

QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
Db 1608 CACACAGAAAGAGAGAAATCCCTACAAAGAGATGTACAGGACATGTGGGTGAACCT 1667

QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
Db 1668 GAGGAGCTGCTTACGACACCTTCCACCGCCAAAGAGCCCGGAGAGCAGCGGAG 1727

QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db 1728 AAGCCCAAGCTCAAGAGATTAATTGATGAGCCACAGAGAGAGCGGCTGTGACGAGTG 1787

QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db 1788 CGCAGAACTGCCGAAACATTGAGACATCTGCATCTCTGTGGAGCCTCAATGTTAC 1847

QY 281 LeuGluHisProLeuPheIleGlyLysMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db 1848 CTGAAACACCCCTCTTCTGTTGAGGATGTCCAAACTGCAGAACTGCTTCTGAG 1907

QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db 1908 TGTGGCTACCACTACAGACAGACGCGCTACAGTCTCACTGACACCATGTGTGGGGG 1967

QY 321 ArgGluValLeuMetCysGlyLysAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db 1968 CGTGAAGTGTCTATGTGGCGAAACAACTGCTCAGGTGTCTTTGCTGTGAGTGTG 2027

QY 341 AspLeuLeuValAGLProGluValAlaGlnAlaIleLysGluLysAspProThrPheCys 360
Db 2028 GACCTTGTGTGGGCGGGGCTGCCAGGAGCCATTAAAGAAACCCCTGGAACCTGC 2087

QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGluAspTrpProSer 380
Db 2088 TACATGTGGGCGCAACAAGGTACTTACGGGCTGCTGGCGGCGAGAGAGATGGCCCTCC 2147

QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
Db 2148 CGGTCTCCAGATTTCTTGTCTAATTAACACGACCAAGAAATTTGACCTCCAAAGTTTAC 2207

QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 2208 CCACTGTCCAGCTGAGAGAGAGAGCCCATCCGGGTGCTCTCTCTTGTATGGAATC 2267

QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db 2268 GCTACAGGGCTCTGTGTGTAAGGACTTGGGCACTTCAAGTGAACGCTACATTGGCTCG 2327

QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db 2328 GAGGTGTGTGAGGATCCATCACTGAGGAGATGTGTGGCGCACCAAGGGAGATCAATGTAC 2387

QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
Db 2388 GTGGGGAGCTCCGAGGCTCACAGAGCATATCCAGAGATGGAGGCCCATTCGATCTCG 2447

QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db 2448 GTGATGGGGGAGAGCTCTGCAATGACTCTTCATCTGCAACCTGCTGCAAGGCGCTC 2507

QY 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db 2508 TACGAGGGGACATGGCGGCTCTTCTTGTAGTTCTACCGCTCTCTCATATATCCGGGCCC 2567

QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
Db 2568 AAGGAGGAGATGATGCCCTTCTTGTGCTCTTGTGAAATGTGTGCTTGGCGCTT 2627

QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominae; Homo.
REFERENCE	1 Totoki, Y., Toyoda, A., Takeda, T., Sakaki, Y., Tanaka, A., Yokoyama, S., Ohara, O., Nagase, T. and F. Kikuno, R.
TITLE	None Title
JOURNAL	Published Only in Database (2005)
AUTHORS	2 (bases 1 to 4476) Totoki, Y., Toyoda, A., Takeda, T., Sakaki, Y., Tanaka, A., Yokoyama, S., Ohara, O., Nagase, T. and F. Kikuno, R.
TITLE	Direct Submission
JOURNAL	Submitted (22-MAR-2005) Osamu Ohara, Kazusa DNA Research Institute, Department of Human Gene Research; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 232-0818, Japan (E-mail:cdnainfo@kazusa.or.jp, URL:http://protein.gsc.riken.go.jp/, Tel:81-438-52-3930, Fax:81-438-52-3931)
COMMENT	This work was supported in part by the National Project on Protein Structural and Functional Analysis, Ministry of Education, Culture , Sports, Science and Technology of Japan. Totoki Y., Toyoda A., , Takeda T., Sakaki Y., Tanaka A., Yokoyama S. RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsumagi, Yokohama 230-0045, Japan. e-mail: aktanaka@postman.riken.go.jp URL: http://protein.gsc.riken.go.jp/ Location/Qualifiers
FEATURES	1. 4476 /organism="Homo sapiens" /mol_type="RNA" /db_xref="taxon:9606" /clone="pj01914" /tissue_type="brain" /note="this clone is also named as hsk003000130 vector:pb1uescriptII SK plus" 1. 4476 /gene=DNA cytosine methyltransferase 3 alpha isoform a variant" 148. 2563 /gene=DNA cytosine methyltransferase 3 alpha isoform a variant" /inference="non-experimental evidence, no additional details recorded" /note="start codon is not identified." /codon_start=1 /product="DNA cytosine methyltransferase 3 alpha isoform a variant" /protein_id="BAB92070.1" /db_xref="GI:62087246" /translation="LSPPKTTTIPSPKSSPPAPRSPSLPQAQMAMSSSGGCDTSS SAERKDRDGEEOEPPRGKERQEPSTTARVGRGRRKRPVPSGDTPDPAVI SKRPSMAODSGASELLPNGLDEKRSRPOEBEGSPAGOGGAPAEDEGAETLPEASR TAVNGCCTPREGGAPAEAGKEQETNISMMKSGRGLRGGLGESSLRQPMPLR FPAGGYYTISKRRKDEMLAWKREBAEKAKYTAGNAVABNGPDSQKVEASPPA VOQPTPASPATYATTPPEPGSDAGDKNATKAGDDEVEYEDGFGTGLVMKTLRGFS WMGRIVSWMTGRSRAAGTRVWMPFGDKRSVCEVKMLSSFCSAPHQATYKQ PMYRKAIYEVLLQVASSRAGLFPVCHDSDESDPAKAEVQNKEIMLAWGAFQPSDPK GLEPPEEKNPYEVYTDMMVEBEAAVAPPPAKKPRKSTAEKPYKSIIDERTER LVEVEKOKGNIEDICISGSLNVTLEHLPFGVGMCKNCKCLECAVYDDGYOSY CTICCGGAEVIMCGNNGNCCRCFVECEVDLVLVGGAAQAIKEDPMNCVACGKHGTGL LRRREMPBSRLQMFANHHDDPFPKYPVPPAERKRIITUSLTDGATGLTVLAKD LGIVQRYTASVEYCEDSTIVGVNRHOGKTIYGDVASVQKHIOEAGFPDLVIGSSPC NDLSIVNPAKGLYEGTGRLFFEFYRLLDHARDKESDDRPFWLFFNVWAMGVSDKRD ISRFLEVONSHLY"
CDS	
gene	
ORIGIN	
Alignment Scores:	
Pred. No.:	0
Score:	3653.00
Percent Similarity:	90.1%
Best Local Similarity:	89.8%
Query Match:	97.2%
DB:	5
	Length: 4476 Matches: 679 Conservative: 2 Mismatches: 8 Indels: 67 Gaps: 1

us-10-623-813-85 (1-688) x AB208833 (1-4476)

QY 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyGluSerGlnIlyValGluGluAlaSer 20
 Db 907 ATGAATGCTGTGGAAGAAAACAGAGGGCCGGGAGTCTCAAGAGGTGAGAGGCCAGC 966

QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 Db 967 CCTCCGCTGTGACGAGCCCACTGAGCCCGCATCCCACTGGGTCTACAGCCCTGAG 1026

QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 Db 1027 CCGGTGGGGTCCCATGCTGGGAGCAAGAAATGCCAAGACGAGCGATGACGAGCCAG 1086

QY 61 TyrGluAspGlyArgGlyPheGlyLeuValThrGlyLysLeuArgGlyPhe 80
 Db 1087 TAGAGAGAGCGCCGGGCTTTGGCATTTGGGAGAGTGTGTGGGAAATCGCGGGCTTC 1146

QY 81 SerTrpTrpProGlyArgGlyLeuAlaSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
 Db 1147 TCTGTGGGCGAGCGCCGATTTGTCTGTGTGATGACGGGCGGAGCCGAGCGCTGAA 1206

QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
 Db 1207 GGCACCCGCTGGGTATGTGTTTGGAGACGGCAAATTCAGTGTGTGTGTGAGAG 1266

QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 Db 1267 CTGATCGCGCTGAGCTGTTTTCAGTGTGCTTCCACGAGGCCAGTCAACAGAGCCC 1326

QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 Db 1327 ATGTACCGCAAGCCATCTAGAGTCTCTGAGTGGCCAGCGCCCGGGGGAGCTG 1386

QY 161 PheProAlaCysHisAspSerAspGlySerAspSerGlyLysAlaValGluAlaGlnAsn 180
 Db 1387 TTCCCGGTGTCCACGACGAGAGTGAAGTGAAGTCACTCCAAAGCCGATGAGGTGCAAG 1446

QY 181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
 Db 1447 AACCCCTGATTGAAATGGCCCTGGGGGCTTCCAGCCTTCTGCTTAAAGGGCTGAG 1506

QY 201 ProProGluGluGlyLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
 Db 1507 CCAACCAAGAAAGAAATCCCTACAAAGATGTAACAGGACATGTGGTGAACCT 1566

QY 221 GlnAlaAlaAlaTyrAlaProProProAlaLysLysProArgLysSerThrThrGlu 240
 Db 1567 GAGGCACTGCTTACGACCACTCCACGCAAAAGCCCGGAGAGACAGACGCGGAG 1626

QY 241 LysProLysValLysGluIleIleAspGluArgThrArgLysLeuValTyrGluVal 260
 Db 1627 AACCCCAAGGTCAAGAGATTATGATGAGGCAAGAGGCGGTGTGTATCAGAGGTG 1686

QY 261 ArgGlnLysCysArgAsnIleGluAspLysCysIleSerCysGlySerLeuAsnValThr 280
 Db 1687 CGGCAAGAGTCCGGAACATTGAGACATCTGATCTCTGTGGAGACCTCAATGTAAAC 1746

QY 281 LeuGlnHisProLeuPheIleGlyGlyMetCysGlnAsnAsnCysLysAsnCysPheLeuGlu 300
 Db 1747 CTGGAACACCCCTCTTCTGTGGAGAAATGTGCCAAAACGTGCAGAACTGCTTCTGAG 1806

QY 301 CysAlaLysGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
 Db 1807 TGTGCGTACAGTACGACGAGCGGTACCACTCACTGACCACTGCTGTGGGGGC 1866

QY 321 ArgGlnValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
 Db 1867 CGTAGAGGTCTCATGTGCGGAACAACAATCTGCTGAGTCTTTTGTGGAGATGTGTG 1926

QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
 Db 1927 GACCTCTTGTGGGGCGGGGGCTGCCAGGCAACATTAAGAAAGCCCTTGAACCTGC 1986

QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
 Db 1987 TACATGTGGGCGACAAAGGTATCTTACGGGCTGTGCGGGCGGAGAGAGATGGCCCTTC 2046

QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
 Db 2047 CCGCTCCAGATGTCTTCCCTTAATAACACGACAGAGAAATTTGACCTCCAAAGTTTAC 2106

QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 Db 2107 CCACCTGTCCCAAGTGAAG 2166

QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAlaAspArgTyrIleAlaSer 440
 Db 2167 GCTACAGGGCTCTGCTGTGTAAGACTTGGGCAATTCAGGTGAGCCGCTACATGTCTTC 2226

QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 Db 2227 GAGGTGTGAGAGACTCCATCACGGTGGCATGTGTGCGCACAGGGAGAAATCATGTAC 2286

QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
 Db 2287 GTCGGGGAGCTCCGCGAGCTCACAGAGACATATCAGAGAGTGGGCCCATTCGATCTG 2346

QY 481 ValIleGlyLysProCysAsnAspLeuSerIleValAlaAspProAlaArgLysGlyLeu 500
 Db 2347 GTGATGGGGGAGAGTCCCTGCAATGACCTTCTCATCTGTAACCTCTCGCAAGGGCTC 2406

QY 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
 Db 2407 TAGAGGGCACTGGCGGCTCTTCTTGAATTCACCGCTCCTGTGATGAGCGGGCCC 2466

QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
 Db 2467 AAGGAGGAGAGATCGCCCTTCTTGTGGAAGTGTGTGTCATGAGGCGT 2526

QY 541 SerAspLysArgAspIleSerArgPheLeu 550
 Db 2527 AGTGACAAAGGAGCACTTCGGCATTTCTCAGAGTCAAAATAGCAACCTGTACTAGTG 2586

QY 550 550
 Db 2587 GCTAGTGGCACTGAATTGAACAGTGTAGAGCAACCAACCATCCAGAGTTCAATTG 2646

QY 550 550
 Db 2647 GACAAATGCTGATCTAGAACGTTCCAGTGAACGTTGTGAATGTGGCTAGGTAAATTC 2706

QY 551 551
 Db 2707 AGCTTTCCTCGAGAGCTGACAGGCTAGCCAGTGTGGCTCTCGAAGAGAAATCCAAACC 2766

QY 554 554
 Db 2767 TGTGATGATTGATGCCAAAGAGTGTCACTGACACAGAGGCCGCTACTTCTGGGTAA 2826

QY 574 574
 Db 2827 CTTTCCCGGATGAAACAGGCCGTTGGCATTCATGTGAATGATTAACCTGAGCTGAGGA 2886

QY 594 594
 Db 2887 GTGTCTGGAGCATGGAGATACCAAGTTCAGCAAAAGTGAAGCAATTAATCAAGGTTC 2946

QY 614 614
 Db 2947 AAATCCATTAAGCAAGGCAAGAGCAAGCATTTTCTGTCTTCAATAATAGAAAGGA 3006

QY 634 634
 Db 3007 CATCTTATGTGTGACATGAAGTGAAGAGGATTTGTGTTCCAGTCCACTATATCTAGCT 3066

Qy	654	1SerAsnMetSerArgLeuAlaArgGlnArgLeuLeuGlyArgSerTrpSerValProVa	674
Db	3067	CTCCAACTATGAGCCGCTTGGGAGGCAAGACTCTGGCCGGTCAATGAGCGTGCAGT	3126
Qy	674	111eArgHisLeuPheAlaProLeuLysGluTyrPheAlaCysVal	689
Db	3127	CATCCGCCACCTCTTCGCTCGCTGAAGAGTATTTTGGGTGTGTG	3172

Search completed: October 10, 2006, 17:31:30
Job time : 10584 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 10, 2006, 14:46:24 : Search time 2084 Seconds
(without alignments)
6093.700 Million cell updates/sec

Title: US-10-623-813-85

Perfect score: 3758
Sequence: 1 MNAVEBNQASGESQKVEAS.....WSVPVIRHLPADLKEYFACV 689

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+gdn.model -DEV=x1h
-O=/abs/ABSWEB.spool/US10623813/runat_10102006_143521_7265/app_query.fasta_1
-DB=Published Applications NA.Main -QFMT=fastcat -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR.SCORE=pct
-THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03h
-USER=US10623813@CCN 1.1.1675 -runat_10102006_143521_7265 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA.Main:

1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11E_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3758	100.0	2318	9	US-10-623-813-83
					Sequence 83, Appl

2	3758	100.0	4192	5	US-09-720-086-1	Sequence 1, Appl1
3	3758	100.0	4192	9	US-10-623-813-1	Sequence 1, Appl1
4	3758	100.0	6157	11	US-10-330-773-276	Sequence 276, App
5	3714	98.8	2371	9	US-10-623-813-84	Sequence 84, Appl1
6	3714	98.8	3005	6	US-10-144-577-2	Sequence 2, Appl1
7	3714	98.8	3005	6	US-10-172-118-132	Sequence 132, App
8	3714	98.8	3005	8	US-10-342-887-132	Sequence 132, App
9	3714	98.8	3009	11	US-10-330-773-279	Sequence 279, App
10	3714	98.8	4293	5	US-09-720-086-3	Sequence 3, Appl1
11	3714	98.8	4293	9	US-10-623-813-3	Sequence 3, Appl1
12	2708	72.1	2338	7	US-10-264-237-953	Sequence 953, App
13	2159.5	57.5	4145	5	US-09-720-086-4	Sequence 4, Appl1
14	2159.5	57.5	4145	5	US-10-144-577-3	Sequence 3, Appl1
15	2159.5	57.5	4145	6	US-10-144-577-38	Sequence 38, Appl
16	2159.5	57.5	4145	7	US-10-172-118-1337	Sequence 1337, Ap
17	2159.5	57.5	4145	8	US-10-342-887-1337	Sequence 1337, Ap
18	2159.5	57.5	4145	9	US-10-623-813-4	Sequence 4, Appl1
19	2159.5	57.5	4195	5	US-09-720-086-2	Sequence 2, Appl1
20	2159.5	57.5	4195	9	US-10-623-813-2	Sequence 2, Appl1
21	1881	50.1	3897	6	US-10-144-577-39	Sequence 39, Appl1
22	1837.5	48.9	95484	11	US-10-330-773-275	Sequence 275, App
23	1722.5	45.8	127722	11	US-10-330-773-278	Sequence 278, App
24	1318	35.1	1060	6	US-10-106-698-926	Sequence 926, App
25	1095.5	29.2	709	3	US-09-969-034-1235	Sequence 1235, Ap
26	1003	26.7	711	10	US-10-450-763-21969	Sequence 21969, A
27	944	25.1	622	12	US-09-764-891-272	Sequence 272, App
28	942	25.1	994	12	US-10-301-480-606665	Sequence 606665,
29	942	25.1	994	12	US-10-301-480-1220074	Sequence 1220074,
30	872.5	23.2	2057	6	US-10-144-577-40	Sequence 40, Appl1
31	785	20.9 <td>2008</td> <td>6</td> <td>US-10-144-577-41</td> <td>Sequence 41, Appl1</td>	2008	6	US-10-144-577-41	Sequence 41, Appl1
32	750	20.0	440	5	US-09-720-086-54	Sequence 54, Appl1
33	750	20.0	440	9	US-10-623-813-54	Sequence 54, Appl1
34	688	18.3 <td>1705</td> <td>8</td> <td>US-10-479-334-35</td> <td>Sequence 35, Appl1</td>	1705	8	US-10-479-334-35	Sequence 35, Appl1
35	685	18.2 <td>2951</td> <td>10</td> <td>US-10-450-763-21970</td> <td>Sequence 21970, A</td>	2951	10	US-10-450-763-21970	Sequence 21970, A
36	644.5	17.2	1670	8	US-10-479-334-21	Sequence 21, Appl1
37	639	17.0	1521	8	US-10-479-334-7	Sequence 7, Appl1
38	636	16.9	522	5	US-09-720-086-61	Sequence 61, Appl1
39	636	16.9	522	9	US-10-623-813-61	Sequence 61, Appl1
40	629	16.7	598	6	US-10-027-633-199022	Sequence 199022,
41	629	16.7	598	6	US-10-027-633-199022	Sequence 199022,
42	629	16.7	598	6	US-10-027-633-199022	Sequence 199022,
43	629	16.7	598	7	US-10-027-633-199023	Sequence 199023,
44	609	16.2	435	5	US-09-720-086-18	Sequence 18, Appl1
45	609	16.2	435	9	US-10-623-813-18	Sequence 18, Appl1

ALIGNMENTS

RESULT 1
US-10-623-813-83
Sequence 83, Application US/10623813
Publication No. US20040234997A1
GENERAL INFORMATION:
APPLICANT: LI, En
APPLICANT: Okano, Masaki
APPLICANT: Xie, Shaoping
APPLICANT: Chen, Taiping
TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Use
FILE REFERENCE: 0609.4560003
CURRENT APPLICATION NUMBER: US/10/623,813
CURRENT FILING DATE: 2003-07-22
PRIOR APPLICATION NUMBER: US 09/720, 086
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: PCT/US99/14373
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 60/090, 906
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: US 60/093, 993
PRIOR FILING DATE: 1998-07-24
NUMBER OF SEQ ID NOS: 119
SOFTWARE: PatentIn version 3.2
SEQ ID NO 83

```

; LENGTH: 2318
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-623-813-83

Alignment Scores:
Pred. No.: 0 Length: 2318
Score: 3758.00 Matches: 689
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0

US-10-623-813-85 (1-689) x US-10-623-813-83 (1-2318)
Qy      1 MetAsnAlaValGluGluAsnGlnAlaSerGlyGluSerGlnValGluGlnAlaSer 20
Db      148 ATCAATGCTGTGGAAGAAACCAAGCCTCTGAGAGTCTCAAGAGTGAAGAGCCAGC 207
Qy      21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db      208 CCTCCTGCTGTGACGACGCCACGGAACCTGCTTCCGACTGTGGCCACACCCCTGAG 267
Qy      41 ProValGlyGlyAspAlaGlyAspIleAsnAlaThrIleValAlaAspAspGluProGlu 60
Db      268 CCAATGGAAGGAGATGCTGGGGACAAAGATGCTCAAGACAGCCGACATGAGCTTGAG 327
Qy      61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValThrGlyIleAsnArgGlyPhe 80
Db      328 TATGAGGATGGCCGGGCTTTGGCATTTGAGAGCTGTGTGGGGAAACTTGGGGCTTC 387
Qy      81 SerTyrTrpProGlyArgIleValSerTyrTrpMetThrGlyIleArgSerAlaAlaGlu 100
Db      388 TCCTGTGGGCAAGCCCAATTTGTCTTGTGGATGACAGCCGAGCCGAGAGCTGAA 447
Qy      101 GlyThrArgThrValMetTrpPheGlyAspGlyIleAspSerValValCysValGluIle 120
Db      448 GGCACCTGGCTGGGTCATGTGTTGGAGATGGCAAGTTCTCAAGTGTGTGTGGAGAG 507
Qy      121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnIleGlnPro 140
Db      508 CTGATCCCGTGAAGCTCTCTCTGAGTGCATTTCCACAGGCCACTTACAAAGCAGCC 567
Qy      141 MetTyrArgIleAlaIleTyrGluValIleGlnValAlaSerSerAlaAlaGlyIleLeu 160
Db      568 ATCTACCGCAAAAGCATCTACGAAGTCTCAGAGTGGCCAGAGCCCTGGCGGAAGCTG 627
Qy      161 PheProAlaCysHisAspSerAspGluSerAspSerGlyIleValAlaGluValGlnAsn 180
Db      628 TTTCAGAGCTTGCATGACATGATGAAGATGAAGTGGCAAGGCTGTGGAATGCGAG 687
Qy      181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProIleGlyLeuGlu 200
Db      688 AACCAAGTATGATGATGGCCCTCGTGGCTTCCAGCCCTCGGCTCTTAAGGGGCTGAG 747
Qy      201 ProProGluGluGluIleAsnProTyrIleGluValIleThrAspMetTrpValGluPro 220
Db      748 CCACCAAGAAAGAAAGAAATCTCTTACAAAGAAATTTACCAAGCAAGTGGGTGAACCT 807
Qy      221 GluAlaAlaIleTyrAlaProProProProAlaIleIleValIleValIleValIle 240
Db      808 GAAGACAGCTGCTTACGCCCAACCCCAACCAAGCAAGAAAGCAAGCAAGCAAGAG 867
Qy      241 LysProIleValIleGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db      868 AAACCTAAGGTCAAGAGATCATTTGATGAGGCAAAAGGAGCGGCTGTGTATGAGGTG 927
Qy      261 ArgGlnIleCysArgAsnIleGluAspIleCysIleSerCysGlySerIleAsnValThr 280
Db      928 CGCCAGAAAGTCAAGAAATCATGAGACATTTGATCTCATGTGGAGCCCTCAAGTCAAC 987
Qy      281 LeuGluIleProLeuPheIleGlyIleCysGlnIleAsnCysIleAsnCysPheLeuGlu 300

```

```

Db      988 CTGAGACACCACTCTTCAATTGGAGCATGTGCCAAGACTGTAAAGACTCTTCTTGAG 1047
Qy      301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      1048 TGTGCTTACAGATATGACACACATGTGGTACCATCTTATGACCACTCTGCTGTGGGGG 1107
Qy      321 ArgGluValLeuMetCysGlyIleAsnAsnIleAsnIleCysArgCysPheCysValGluCysVal 340
Db      1108 CCGAAGTGTCTCATGTGTGGAAACAACCTGTGAGGTCTTTGTGTCCAGTGTGTG 1167
Qy      341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProThrAsnCys 360
Db      1168 GATCTCTTGTGGGGCAAGAGCTGTCTGAGGACATTAAAGAAAGACCCCTGGAACCTGC 1227
Qy      361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGluAspTrpProSer 380
Db      1228 TACATGTGGGGCATTAAGGACCTATGTGGCTGTCCGAAGACGGGAACATGGCTTGT 1287
Qy      381 ArgLeuGlnMetPhePheAlaAsnAsnIleAspGlnIlePheAspProProIleValTyr 400
Db      1288 CAGCTCCAGATGTTCTTTCCAATTAACATGACCAAGAAATTTGACCCCAAGATTAC 1347
Qy      401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      1348 CCACTGTGCCAGCTGAGAAAGAAAGCCCATCCGGCTGTCTCTTGTGATGGGATT 1407
Qy      421 AlathrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      1408 GCTACAGGGCTCTCTGTGTGAAGACCTGGGCATTCAGTGAACCGCTACATGCTCC 1467
Qy      441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnIleLysIleMetTyr 460
Db      1468 GAGGTGTGAGGACTCTCCATCACGGTGGCATGTGTGGCACCAGGAAAGATCATGTAC 1527
Qy      461 ValGlyAspValArgSerValThrGlnIleHisIleGlnIleTrpGlyProPheAspLeu 480
Db      1528 GTCCGGGAGCTCCGACGCTCACACAGAGCATATTCAGAGATGGGGCCCATTCGACCTG 1587
Qy      481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgIleGlyLeu 500
Db      1588 GTGATTTGGAGGAGTCCCTGCAATGACCTCTCATTTGAACCTTCCGCAAGGACATT 1647
Qy      501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db      1648 TATGAGGATCTGGCGGCTCTTCTTGAAGTCTACCGCTCTCTGATGATGGCGGCC 1707
Qy      521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaIleMetGlyVal 540
Db      1708 AAGGAGGAGATGATGCGCCCTTCTTCTGTGAGAAATGTGTGGCCCATGGCGCTT 1767
Qy      541 SerAspLysArgAspIleSerArgPheLeuGluIleAsnProValMetIleAspAlaLys 560
Db      1768 AGTGACAAAGGAGAACTTCGGATTTCTTGAGTCTTAACCCCGTGAATGACGCCAAG 1827
Qy      561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyIleMetAsnArg 580
Db      1828 GAAGTGTCTGTGCACACAGGCGCTTACTTGTGGGTAACCTTCTGTGCAATGAACAG 1887
Qy      581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGluCysLeuGluHisGlyArg 600
Db      1888 CTTTGGCATCCACTGTGAATGATTAAGCTGAGCTGCAAGAGTGTCTGAGACCGCAAG 1947
Qy      601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnIle 620
Db      1948 ATAGCCAAAGTTCAAGAAATGAGGACCATTTACCAAGGCTCAAACTTTAAAGCAGGGC 2007
Qy      621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db      2008 AAAGACCAACATTTCCCGCTCTTCAATGAACGAGAAAGAGACATCTGTGTGTCACTGAA 2067
Qy      641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2068 ATGAAAGAGGTGTTGTGCTTCCCGTCCACTACACAGAGCTCTCAACATGAGCCGCTTG 2127

```

QY 661 AlaArgGlnArgLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
 DB 2128 GCGAGGCGAGACACTGCTGGCCGATCTGAGAGCTGCCGATCATCCGCCACTCTTCGCT 2187
 QY 681 ProLeuIysGluTrpPheAlaCysVal 689
 DB 2188 CCGCTGAAGGAATATTTTGTGTTGTG 2214
 RESULT 2
 US-09-720-086-1
 ; Sequence 1, Application US/09720086
 ; Publication No. US20060084053A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Li, En
 ; APPLICANT: Okano, Masaki
 ; APPLICANT: Xie, Shaoping
 ; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,
 ; TITLE OF INVENTION: Polypeptides & Uses Thereof
 ; FILE REFERENCE: 0609, 4560002
 ; CURRENT APPLICATION NUMBER: US/09/720, 086
 ; CURRENT FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US99/14373
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: 60/090,906
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/093,993
 ; PRIOR FILING DATE: 1998-07-24
 ; NUMBER OF SEQ ID NOS: 82
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 4192
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (4161)..(4161)
 ; OTHER INFORMATION: May be any nucleic acid
 US-09-720-086-1
 Alignment Scores:
 Pred. No.: 0 Length: 4192
 Score: 3758.00 Matches: 689
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 5 Gaps: 0
 US-10-623-813-85 (1-689) x US-09-720-086-1 (1-4192)
 QY 1 MetaAsnAlaValGluGluAsnGlnAlaSerGlyGluSerGlnIysValGluGluAlaSer 20
 DB 874 ATCAATGCTGTGGAAGAAGCAAGGCTCTGAGAGCTCAGAAAGTGAGGAGGCGCAGC 933
 QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrTrpProGlu 40
 DB 934 CCTCTGCTGTGACGACGCCACGACCTGCTTCTCCGACTGCTGGCCACCACTCCCTGAG 993
 QY 41 ProValGlyGlyAspAlaGlyAspIysAsnAlaThrIysAlaAlaAspAspGluProGlu 60
 DB 994 CCAAGTGAAGGAGGATGCTGGGGAACAAGATGCTACCAAGACACCGACGATGAGCTGAG 1053
 QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyIysLeuAlaArgGlyPhe 80
 DB 1054 TATGAGAGATGCGCGGCTTTGGCATTGAGAGACTGCTGGGGAACCTTCGGGGCTTC 1113
 QY 81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
 DB 1114 TCTGTGTGGCAAGCGCAATGTGTCTGTGTGATGACAGCGCGAGCCGACGACTGAA 1173
 QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyIysPheSerValValCysValGluIys 120
 DB 1174 GGCACTGCTGCTCATGTGTGTCGAGATGCAAGATTCTCAGTGTGTGTGTGAGAG 1233

QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTrpAsnIysGlnPro 140
 DB 1234 CTATAGCCGCTGAGCTTCTTCTGACATTCATCCAGGCCCACTTACACAAAGCAAGCCC 1293
 QY 141 MetTyrArgIysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyIysLeu 160
 DB 1294 ATGTACCGAAGACCATTCATCAAGAGTCTTCAGAGTGGCCAGCGCGGTGCGGGAAGCTG 1353
 QY 161 PheProAlaCysHisAspSerAspGluSerSerGlyIysAlaValGluValGlnAsn 180
 DB 1354 TTTCACCTTGGCCATATACAGATGTGAAGTGAACAGTGGCAAGGCTGTGAAGTGCAGAAC 1413
 QY 181 LysGluMetIleGluTrpAlaLeuGlyIysPheGluProSerGlyProIysGlyLeuGlu 200
 DB 1414 AACCAATGATTAATGAGGCCCTCGGTGCTTCAAGCTCTCGAGTCTTAAGGCTGTGAG 1473
 QY 201 ProProGluGluGluIysAsnProTyrIysGluValTyrThrAspMetTrpValGluPro 220
 DB 1474 CCACGAAAGAAAGAAAGAAATCCTTCAAGAAAGTTTACCCGACATGTGTGTGAGCCT 1533
 QY 221 GluAlaAlaAlaTyrAlaProProProProAlaIysIysProArgIysSerThrTrpGlu 240
 DB 1534 GAAGCAGCTGCTTACCCGCCACCCACCCACGCAAGAAACCCAGAAAGACACAAACAGAG 1593
 QY 241 LysProIysValIysGluIleIleAspGluArgTrpArgGluArgLeuValTyrGluVal 260
 DB 1594 AAACCTTAAGGTCAAGAGATCATTGATGAGCGCAAGAGGAGCGCTGTGTGTGAGGTG 1653
 QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 1654 CGCCAAAGTGAAGAAATCCAGAGCATTTGTATCTCATGTGTGAGCGCTCAATGTACACC 1713
 QY 281 LeuGluHisProLeuPheIleGlyIysMetCysGlnAsnCysIysAsnCysPheLeuGlu 300
 DB 1714 CTGGACACCCCACTTCTTATGGAGCATGTGCACAACTGTAAAGACTTCTTGTGGAG 1773
 QY 301 CysAlaIleTyrGlnTrpAspAspGlyIysGlnSerTyrCysThrIleCysCysGlyIys 320
 DB 1774 TGTGCTTACAGATGACACAGATGGGTACAGTCTTATTTGACACTGTGCTGTGGGGG 1833
 QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 340
 DB 1834 CCGAAGTCTCATGTGTGTGGAACAACAACCTGTGAGGTGCTTTGTGTGCGAGTGTGTG 1893
 QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleIysGluAspProTrpAsnCys 360
 DB 1894 GATCTCTGTGTGGGCGCAAGAGCTGCTCAGGACGACATTAAAGAAACCCCTGGAACCTGC 1953
 QY 361 TyrMetCysGlyHisIysGlyThrTyrGlyLeuLeuArgArgGluAspTrpProSer 380
 DB 1954 TACATGTGGGCAATAGGGAACCTTAGGGCTGTGCGAAGAGGGGAAGACTGGCTTCT 2013
 QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProIysValTyr 400
 DB 2014 CCACTCAAGATGTTCTTCCATTAACATGACCAAGAAATTTGACCCCAAGAGTTTAC 2073
 QY 401 ProProValProAlaGluIysArgIysProIleArgValLeuSerLeuPheAspGlyIle 420
 DB 2074 CCACTGTGCGCAGCTGAGAAAGGAAAGCCACCTCGGTGCTCTCTCTTGAATGGGAAT 2133
 QY 421 AlaThrGlyLeuLeuValLeuIysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
 DB 2134 GCTACAGGCTCTGTGTGCTGAAGGACTGGGCAATCCAGTGAACCGCTACATTGCTTCC 2193
 QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyIysIleMetTyr 460
 DB 2194 GAGGTGTGTGAAGACTCCATCAACGTTGGGCAATGTGTGCGCAACAGGAAAGATCATGTAC 2253
 QY 461 ValGlyAspValArgSerValThrGlnIysHisIleGlnGluTrpGlyProPheAspLeu 480
 DB 2254 GTGCGGAGCGTCCGAGGCTCAACAGAAAGCATATTCAGGAATGGGGCCCACTTGCACCTG 2313

```

QY 481 valilleglyserprocyasnaaspleuSerileValaenproalaaarglysglyleu 500
DB 2314 GTGATTGGAGGAGTCCCTGCAATGACCTTCATGTCACCTGCGCCCAAGGACTT 2373
QY 501 TTTGGLGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
DB 2374 TATGAGGGTACTGGCCGCTCTTCTTGAGTTCTACCCGCTCTCGCATGATGCCGCCC 2433
QY 521 LysGluGlyAspAspArgProPhePheTyrLeuPheGluAsnValAlaIleMetGlyVal 540
DB 2434 AAGGAGGACATGATGCGCCCTTCTTGCTTGTAGATGATGGTGGCCATGGCGCTT 2493
QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
DB 2494 AGTGACAAAGGAGGACATCTGCGCATTTCTTGAGTCACCCCGCATGATGACGCCAA 2553
QY 561 GluValSerAlaIleHisArgAlaArgTyrPheTyrPheGlyAsnLeuProGlyMetAsnArg 580
DB 2554 GAAGTGTCTCTGACACAGGGCCGCTTACTTCTGGGGTAACTTCTGGCATGAAACAG 2613
QY 581 ProLeuAlaSerThrValaAsnAspLysLeuGluLeuGlnGlyCysLeuGluHisGlyArg 600
DB 2614 CCTTGGCATCCACTGGAATGATTAAGCTGAGCTGCAAGATGCTCGAGACAGCGAGA 2673
QY 601 IleAlaLysPheSerLysValArgThrIleThrTyrArgSerAsnSerIleLysGlnGly 620
DB 2674 ATAGCCAAAGTTCAGCAAGAGGACCAATTAACCCAGGTCAACTTATAAGCAGGGC 2733
QY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTyrCysThrGlu 640
DB 2734 AAGGACCAAGATTTCCCGCTTTCATGACAGAGAGAGACATCTCTGTGTGACAGAA 2793
QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
DB 2794 ATGAAAGGGTGTGGCTTCCCGCTTCCATGACAGAGAGAGACATCTCTGAGCGCTTG 2853
QY 661 AlaArgGluArgLeuLeuGlyArgSerTyrSerValProValIleArgHisLeuPheAla 680
DB 2854 GCGAGCGAGACTGCTGGGCGCGCATCGTGGAGCGTGGCATCCGCACTTCTTGCT 2913
QY 681 ProLeuLysGluTyrPheAlaCysVal 689
DB 2914 CCGCTGAGAGAAATATTTTGTGTGTG 2940

RESULT 3
US-10-623-813-1
; Sequence 1, Application US/10623813
; Publication No. US20040234997A1
; GENERAL INFORMATION:
; APPLICANT: Li, En
; APPLICANT: Okano, Masaki
; APPLICANT: Xie, Shaoping
; APPLICANT: Chen, Taiping
; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
; TITLE OF INVENTION: Therocf
; FILE REFERENCE: 0609,4560003
; CURRENT APPLICATION NUMBER: US/10/623,813
; PRIOR APPLICATION NUMBER: 2003-07-22
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/US99/14373
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,906
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 60/093,993
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 4192
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:

```

```

; NAME/KEY: misc feature
; LOCATION: (4161)..(4161)
; OTHER INFORMATION: n is a, c, g, or t
US-10-623-813-1

Alignment Scores:
Pred. No.: 0 Length: 4192
Score: 3758.00 Matches: 689
Percent Similarity: 100.04 Conservative: 0
Best Local Similarity: 100.04 Mismatches: 0
Query Match: 100.04 Indels: 0
DB: 9 Gaps: 0

US-10-623-813-85 (1-689) x US-10-623-813-1 (1-4192)
QY 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyGlySerGlnLysValGluGluAlaSer 20
DB 874 ATGAATGCTGTGGAAAGAAACCAAGCCCTCTGGAGATCTCAGAACGTTGGAGGAGCCAGC 933
QY 21 ProProAlaValGluGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
DB 934 CTTCTGCTGTGACAGAGCCCAAGACCTGCTTCTCCAGCTGTGGCCACACCTCTGAG 993
QY 41 ProValGlyLysAspAlaGlyAspLysAsnAlaThrLysAlaIleAspAspGluProGlu 60
DB 994 CCAAGTGAAGGGATGCTGGGGACAAAGATGCTACCAAGACCCAGCATGAGCTGAG 1053
QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValITrGlyLysLeuArgLysPhe 80
DB 1054 TATGAGAGATGGCCGGGGCTTTGGCATTTGAGAGCTGTGTGGGGAACCTTGGGGCTTC 1113
QY 81 SerTyrTrpProGlyArgIleValSerTyrTrpMetThrGlyArgSerArgAlaIleGlu 100
DB 1114 TCCTGATGGCAGAGCCGAATGTGTCTTGTGTGATACAGGCCGAGACGACGACTGAA 1173
QY 101 GlyThrArgTyrValMetTyrPheGlyAspGlyLysPheSerValCysValGluLys 120
DB 1174 GGCATCGCTGGGTATGATGTGTTGCGAGATGGCAATTCACATGTGTGTGTGGAGAA 1233
QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
DB 1234 CTCATCCCGCTGAGCTCTTCTGTGACGTGATTCACAGGCCCACTTAACAAGAGCCCTC 1293
QY 141 MetTyrArgLysAlaIleTyrGluValLeuGluValAlaSerSerArgAlaGlyLysLeu 160
DB 1294 ATGTACCGAAAGCCATCTACAGATCTTCAGAGTCCAGACGACGCGTGGGGAAGCTG 1353
QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluAlaAsn 180
DB 1354 TTTCAGCTTGGCATGACATGATGAAGTGAAGTGAAGTGAAGCTGTGCAAGTGCAGAAC 1413
QY 181 LysGluMetIleGluITrPalaLeuGlyPheGlnProSerGlyProLysGlyLeuGlu 200
DB 1414 AAGCAGATGATTTGAAGGGCCCTCGGTGGCTTCACACCTCGGGTCTTAAGGGCTGGAG 1473
QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTyrValGluPro 220
DB 1474 CCAACGAAAGAGAGAGAGATCTTCAAGAGAGTTTACACCGACATGTGGTGGAGCCT 1533
QY 221 GluAlaIleAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
DB 1534 GAAGAGCTGCTTACCCCAACCCCAACCAAGCAAGAAACCCAGAAAGAGACACAGAG 1593
QY 241 LysProLysValLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 260
DB 1594 AAACCTTAAGTCAAGAGATCATTAATGAGCGCAAGAGGAGCGCTGGTGTATGAGGTG 1653
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
DB 1654 CGCCAGAGTGCAGAAATCATCGAGGACATTTGTATCTCATGTGGAGCCCAATGTACCC 1713
QY 281 LeuGluHisPheLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300

```

```

Db      1714 CTGGAGACCACTTCTTCAATGGAGCATGTGCGCAAGACTGTAAAGACTCTTGAGG 1773
Qy      301 CysAlaTyrgLThrAspAspSGlyYrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      1774 TGTGCTTACAGATATACAGACGATGGGTACCACTCTATTGACCACTGTGTGGGGG 1833
Qy      321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGlyCysVal 340
Db      1834 CCGTAAGTCTCATGTGTGGGAACAACACTGCTGACAGGCTCTTTGTGTGCACTGTGTG 1893
Qy      341 AspleuLeuValGlyProGlyAlaAlaGlnAlaAlaIleGlyGluAspProTyrPancCys 360
Db      1894 GATCTCTGTGGGGCCAGAGAGCTGCTCAAGGACCATTAAGAGAACCCCTTGAAGCTGC 1953
Qy      361 TyrMetCysGlyNHisGlyGlyThrTyrGlyLeuLeuArgArgGlyLysPrtProSer 380
Db      1954 TACATGTGGGGCATTAAGGCACTTAATGGCTGCTCGAAGAGGGAGAGCTGGCTTCT 2013
Qy      381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProTyrValTyr 400
Db      2014 CCACTCCAGATGTTCTTTCGCAATACCATGACCAAGAAATTGACCCCAAGATTAC 2073
Qy      401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      2074 CCACCTGTCCAGCTGAGAGAGAGAGAGCCCATCCGCTGCTGCTCTTGTATGGGATT 2133
Qy      421 AlaThrGlyLeuLeuValLeuLysAspleuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      2134 GCTACAGGGCTCTGCTGTGTGAAGACCTGGGCATCAAGTGAACCGCATTCCTCC 2193
Qy      441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db      2194 GAGGTGTGTGAAGACTCCATCAAGGTGGGCTGTGTGGGCAACAGGAAAGATCATCTAC 2253
Qy      461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrPrtPheAspleu 480
Db      2254 GTCGGGGAGCTCCGACGCTCACACAGAACATATACAGAGAGGGGCCCATTCGACCTG 2313
Qy      481 ValIleGlyGlySerProCysAsnAspleuSerIleValAsnProAlaArgLysGlyLeu 500
Db      2314 GTGATTGGAGGAGCTCCCTGCAATGACCTTCCATTGTCAACCTCCCGCAAGGAGATT 2373
Qy      501 TyrGlnGlyThrGlyValArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db      2374 TATGAGGAGTACGTGGCGCTCTTCTTGAGTTCTACCGCTCTTGATATGGCGGCGCC 2433
Qy      521 LysGlnGlyAspAspArgPrtPhePheTyrPleuPheGluAsnValAlaIleMetGlyVal 540
Db      2434 AAGGAGGAGATGATGCGCCCTTCTTGTGGCTTGTGAGATGTGTGGCCATGGCGATT 2493
Qy      541 SerAspLysArgAspIleSerArgPheLeuGlnSerAsnProValMetIleAspAlaLys 560
Db      2494 AGTGCACAAAGGAGCACTTCGCGATTTCTTGAGTCTTAACCCCGTGAATGATGCGCAAA 2553
Qy      561 GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArg 580
Db      2554 GAAGTCTCTGCTGCACACAGGCGCTTACTTGTGGGTAACTTCTTGCAAGAACAG 2613
Qy      581 ProLeuAlaSerThrValAsnAspLysLeuGlnGlnLysLeuGlnHisGlyArg 600
Db      2614 CCTTGGGATCCACTGTGATGATTAAGTGAAGCTCAAGAGTGTGTGGAGCGGCACA 2673
Qy      601 IleAlaLysPheSerLysValArgThrIleThrArgSerAsnSerIleLysGlnGly 620
Db      2674 ATAGCCAACTTCAGCAAACTGAGACCATTAACACAGGCTCAAACTCTTAAGAGCGGC 2733
Qy      621 LysAspGlnHisPheProValPheMetAsnGlnLysGluAspIleLeuTyrPrtGln 640
Db      2734 AAAGACCACTTTCCCGCTTTCAGAACGAGAGAGAGACATCTGTGTGTCACTGA 2793
Qy      641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2794 ATGGAAAGGAGTGTGTGGCTTCCCGTCCACTACACAGAGCTCTCAACATGAGCCGCTTG 2853

```

```

Qy      661 AlaArgGlnArgLeuLeuGlyArgSerTyrSerValProValIleArgHisLeuPheAla 680
Db      2854 GCGAGGCAAGACTCTGTGGCGCGCATGTGGAGCGTGGGTATCGCCACCTTCCGCT 2913
Qy      681 ProLeuLysGluTyrPheAlaCysVal 689
Db      2914 CCGCTGAAGGAATATTTGCTTGTGTG 2940

RESULT 4
US-10-330-773-276
; Sequence 276, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 276
; LENGTH: 6157
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-330-773-276

Alignment Scores:
Pred. No.: 0 Length: 6157
Score: 3758.00 Matches: 689
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 11 Gaps: 0

US-10-623-813-85 (1-689) x US-10-330-773-276 (1-6157)
Qy      1 MetAlaAlaValAlaGlnGluAsnGlnAlaSerGlyGlnSerGlnLysValGlnGluAlaSer 20
Db      866 ATGAATGCTGTGGAAGAGAACAGAGCCCTTGTGAGAGTCTCAAGAGTGAAGAGGCGCAC 925
Qy      21 ProProAlaValAlaGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db      926 CCTCGCTGTGCAGACAGCCAGCCACTCTTCTCCAGCTGTGGCCACACCCCTGAG 985
Qy      41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
Db      986 CCAAGTAGGAGGAGATGCTGGGACAAAGATGCTACCAAGCCGACGATGAGCTTGAG 1045
Qy      61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80
Db      1046 TATGAGATGGCGGGGCTTGTGCAATTGGAGCTGTGTGGGGAACCTTCGGGCTTC 1105
Qy      81 SerTyrTyrProGlyArgIleValSerTyrTyrMetThrGlyArgSerArgAlaGln 100
Db      1106 TCTGTGTGCGCAGGCGCAATGTGTCTGTGATGACAGCGCGGAGCCGAGACGCTGAA 1165
Qy      101 GlyThrArgTyrValMetTyrPheGlyAspGlyLysPheSerValValCysValGlyLys 120
Db      1166 GGCATCGCTGTGGTCAATGTGTTCGAGATGGCAAGTTCTCAGTGTGTGTGGGAAG 1225
Qy      121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db      1226 CTCATGCGCTAGCTCTTCTGCACTGATCTTCCACAGGCCACCTTCAACAAGAGGCC 1285
Qy      141 MetTyrArgLysAlaIleTyrGlnValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db      1286 ATGTACCGGAAGCCATCTACAGAGTCTCCAGGTGGCAGAGGCGGTCCGGGAAGCTG 1345
Qy      161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsn 180
Db      1346 TTTCAGCTTGCATGACAGTATGAAGTGAACGTGGCAAGGCTGTGGAGTGCAGAAC 1405

```

QY 181 LysGlnMetIleGlnTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGln 200
 DB 1406 AACGACATGATTAATGAGCCCTCGGTGCTTCACAGCCCTCGGATCTTAAGAGCCCTGGAG 1465
 QY 201 ProProGlnGlnGlnLysAsnProTyrLysGlnValTyrThrAspMetTrpValGlnPro 220
 DB 1466 CCACCAAGAAAGAAAGAAATCCTTACAGAAAGTTTACACCAATGTTGGTGGAGCCT 1525
 QY 221 GlnAlaAlaIleTyrAlaProProProProAlaLysLysProAlaGlySerThrThrGln 240
 DB 1526 GAAGCAGCTCTTAACGCCCAACCCCAAGCAAGAAACCAAGAAAGCAACAAGAG 1585
 QY 241 LysProLysValLysGlnIleIleAspGlnLysThrArgGlnLysLeuValTyrGlnVal 260
 DB 1586 AACCTTAAGTCAAGGAGCATTTGATGAGCGCAAGGAGCGGCTGGTGTGAGGTG 1645
 QY 261 ArgGlnLysPheArgAsnIleGlnAspIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 1646 CGCCAGAAAGTCAAGAAACATCGAGACATTTGATCTCATGTGGAGCCTCAAAATGTCAC 1705
 QY 281 LeuGlnIleProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGln 300
 DB 1706 CTGGACACCCCATCTTCACTTGGAGGCAATGCCAAGAACTGAAGAATCTCTTGGAG 1765
 QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
 DB 1766 TGTGCTTACAGATATGACGACGATGGGTACAGCTCCATTTGACCACTCTGCTGGGGGG 1825
 QY 321 ArgGlnValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGlnCysVal 340
 DB 1826 CGTGAAGTGTCAATGTGTGGGAAACAACACTGCTGCGAGGTCTTTGTGTGAGTGTG 1885
 QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGlnAspProTyrPasnCys 360
 DB 1886 GATCTTGTGGGGCCAGAGAGCTGCTCAAGCAAGCACTTAAGAAAGCCCTGGAAGCTGC 1945
 QY 361 TyrMetCysGlyIleLysGlyTyrTyrGlyLeuLeuArgArgGlnAspTrpProSer 380
 DB 1946 TACATGTGGCGGCAATAGGCGACCTTAATGGCTGTGTGGAGAGCGGGAAGACTGGCTTCT 2005
 QY 381 ArgLeuGlnMerPhePheAlaAsnAsnHisAspGlnGlnLysPheAspProLysValTyr 400
 DB 2006 CGACTCCAGATGTTCTTGGCAATACCAAGCAAGCAATTTGACCCCAAGAGTTTAC 2065
 QY 401 ProProValProAlaGlnLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 DB 2066 CCACTGTGCTGCACTGGAAGAGAGAGCCCATCCGCTGCTCTTCTTGAATGGATT 2125
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
 DB 2126 GCTACAGGGCTCTGTGGTGAAGGACCTGGGCACTCCAACTGAGCCCTTCACTTGGCTCC 2185
 QY 441 GlnValCysGlnLysPheIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 DB 2186 GAGGTGTGTAAGGACTCATACGCTGAGTGGCATGTGGGCAAGGGAAGATCATATTAC 2245
 QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnLysTrpGlyProPheAspLeu 480
 DB 2246 GTGGGGGACCTCCGCAAGCTCAACAAGACATATCCAGAGGTGGGCCCATTCGACTG 2305
 QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2306 GTGATTGGAGGAGTCCCTCAATGACCTCTCCCTTTCACACCTTGGCCCAAGGACCTT 2365
 QY 501 TyrGlnGlyTyrArgLysLeuPhePheGlnPheTyrArgLeuLeuHisAspAlaArgPro 520
 DB 2366 TATGAGGTAATCTGCGCCCTCTTCTTGAAGTCTTACCGCTCTTCGACATGATGCGGGCCC 2425
 QY 521 LysGlnGlyAspAspArgProPhePheTrpLeuPheGlnLysValValAlaMetGlyVal 540
 DB 2426 AAGGAGGAGATGATCCGCCCTTCTTGTGCTCTTTAGAAATGTGTGGCCATGCGGCTT 2485

QY 541 SerAspLysArgAspIleSerArgPheLeuGlnLysAsnProValMetIleAspAlaLys 560
 DB 2486 AGTGAACAAGAGGAGATTCGCGATTTCTTGAAGTCTTAACCCCGTATGATTAAGCCCAA 2545
 QY 561 GlnValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
 DB 2546 GAAAGTGTGCTGCACACAGGCGCCGTTACTTCTGGGGTAACTTCTCGCATGAACAG 2605
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGlnLysCysLeuGlnIleGlyArg 600
 DB 2606 CTTTGGCATCCACTGTGATGATTAAGCTGAGCTGCACAAAGATGTCTTGAGAGACGGCAGA 2665
 QY 601 IleAlaLysPheSerLysValArgThrIleThrTrpArgSerAsnSerIleLysGlnGly 620
 DB 2666 ATAGCCAAAGTCAAGCAAGTGAAGACCATTTACCAAGCTCAAACTTAAAGCAGGCG 2725
 QY 621 LysAspGlnHisPheProValPheMetAsnGlnLysGlnAspIleLeuTrpCysThrGln 640
 DB 2726 AAAGACAGCATTTCCCGCTTTCATGAACGAAAGAGACATCCTGTGTGCACTGAA 2785
 QY 641 MetGlnArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 DB 2786 ATGAAAGGCTGTGGCTTCCCGCTCACTACAGAGAGTCTCCAAACATGAGCGCTTG 2845
 QY 661 AlaArgGlnArgLeuGlnLysArgSerTrpSerValProValIleArgHisLeuPheAla 680
 DB 2846 GGAAGGCAAGACTGTGGCGCGATCGTGAGACGTCGCCGTATCCGCCACTTCCGCT 2905
 QY 681 ProLeuLysGlnTyrPheAlaCysVal 689
 DB 2906 CCGCTGAAGAAATATTTCTTGTGTG 2932

RESULT 5
 US-10-623-813-84
 ; Sequence 84, Application US/10623813
 ; Publication No. US2004023497A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Id. En
 ; APPLICANT: Okano, Masaki
 ; APPLICANT: Xie, Shaoping
 ; APPLICANT: Chen, Taiping
 ; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Use
 ; TITLE OF INVENTION: Therapeutic
 ; FILE REFERENCE: 0609 4560003
 ; CURRENT APPLICATION NUMBER: US/10/623, 813
 ; CURRENT FILING DATE: 2003-07-22
 ; PRIOR APPLICATION NUMBER: US 09/720, 086
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: PCT/US99/14373
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: US 60/090, 906
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: US 60/093, 993
 ; PRIOR FILING DATE: 1998-07-24
 ; NUMBER OF SEQ ID NOS: 119
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 84
 ; LENGTH: 2371
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-623-813-84
 Alignment Scores:
 Pred. No.: 0 Length: 2371
 Score: 3714.00 Matches: 679
 Percent Similarity: 98.8% Conservative: 2
 Best Local Similarity: 98.5% Mismatches: 8
 Query Match: 98.8% Indels: 0
 DB: 9 Gaps: 0
 US-10-623-813-85 (1-689) x US-10-623-813-84 (1-2371)
 QY 1 MetAsnAlaValGlnGlnLysGlnLysSerGlyLysSerGlnLysValGlnGlnLysSer 20

```

Db      217 ATAAATGCTGTGGAAGAAAACAAGGGCCCGGGAGCTCTCAAGAGTGGAGAGGCCAGC 276
Qy      21  ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db      277 CCTCTGCTGTGAGCAGCCCACTGACCCCGCACTCCCTGAGTGTGCTACCAAGCCCTGAG 336
Qy      41  ProValGlyGlyAspAlaGlyAspIleAspAlaThrIleValAlaAlaAspAspGluProGlu 60
Db      337 CCGGTGGGTCCCATGCTGGGGACAGAAATGCCAAGACAGCCAGATGACGAGCCAGAG 396
Qy      61  TyrGluAspGlyArgGlyPheGlyIleGlyGlyIleValIleProGlyIleAspArgGlyPhe 80
Db      397 TACGAGACGCGCGCGGCTTTGGCATTGGGAGCTGTGTGGGGGAAACCTGGGGCTTC 456
Qy      81  SerTyrTrpProGlyArgIleValSerTyrTrpMetThrGlyArgSerArgAlaAlaGlu 100
Db      457 TCTGTGTGGCCAGAGCGCGCATTTGTCTGTGTGATGACGGGCGGAGCCGACAGCTGAA 516
Qy      101 GlyThrArgTrpValMetTrpPheGlyAspGlyIlePheSerValValCysValGlyIle 120
Db      517 GGCACCCGCTGGGTCAATGTGTTCGAGAGAGCGCAAAATCTCAAGTGTGTGTGAGAG 576
Qy      121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnIleGlyPro 140
Db      577 CTGATGCCCTGAGCTCGTTTTCAGTGCCTTCCACAGGCCACGTRCAACAGCAGCC 636
Qy      141 MetTyrArgIleValIleTyrGluValIleGluValAlaSerSerArgAlaGlyIleVal 160
Db      637 ATGTACCGCAAGACCATCTACGAGGTCTCTCAAGTGGCCAGCGCGCGGGGAAAGCTG 696
Qy      161 PheProAlaCysHisAspSerAspGluSerAspSerGlyValAlaAlaGluValGlnAsn 180
Db      697 TTCCTGGTGTGCACACAGCGATGAGTACGACCTCCAGGCGCTGTGAGGTGACGAAC 756
Qy      181 LysGlnMetIleGluTrpAlaIleGlyIleGlyPheGlnProSerGlyProIleGlyLeu 200
Db      757 AAGCCCATATTTAAATGGGCCCTCGGGGGCTTCACCTTCCTGCGCTTAAGGCTCGAG 816
Qy      201 ProProGluGluGluIleLysAsnProTyrLysGluValIleTyrThrAspMetTrpValGlu 220
Db      817 CCACCGAAGAGAGAGAAATCCCTCAAGAAAGTACACGAGCATGTGTGGTGAACCT 876
Qy      221 GluAlaAlaAlaIleTyrIleAspProProProAlaIleLysProArgLysSerThrTrpGlu 240
Db      877 GAGGACAGCTGCTACGACCACTCCACAGCCAAAGCCCGGAGAGACACAGCGGAG 936
Qy      241 LysProLysValIleGluIleIleAspGluArgThrArgGluArgGluValIleGlyIleVal 260
Db      937 AAGCCCAAGGTCAAGAGATTAATGATGAGCCGACAGAGAGCGGCTGTGTACAGGTG 996
Qy      261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db      997 CGGCAAGAGTGC CGGAACATTGAGACATCTGCATCTCTGTGGAGCTCAATGTTACC 1056
Qy      281 LeuGluHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db      1057 CTGGAAACACCCCTCTTCTGTGAGGAATGTCCAAACCTGCAAGACATGCTTTCGAG 1116
Qy      301 CysAlaTyrGlnIleTyrAspAspArgIleTyrGlnSerTyrCysThrIleCysCysGlyIle 320
Db      1117 TCTGCGTACCACTACGACACGACGCGCTACGCTCTACTGACCACTCTGTGTGGGGG 1176
Qy      321 ArgGluValIleuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db      1177 CGTGAAGTCTCATGTGTGCGGAACAACATCTGCTGAGGTGCTTGTGCTGTGAGTGTG 1236
Qy      341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProThrAsnCys 360
Db      1237 GACCTCTTGTGTGGGGCGGGGGCTGCGACGAGCCATTAAAGAAACCCCTGGAACTGC 1296
Qy      361 TyrMetCysGlyHisLysGlyThrTyrGlyIleLeuAspArgArgGluAspTrpProSer 380

```

```

Db      1297 TACATGTGGGGACACAAGGTACTTACGGGCTGCTGGGGCGGAGAGGACTGGCCCTCC 1356
Qy      381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValIle 400
Db      1357 CGGCTCCAGATGTTCTTCCCTAATAACACAGCACAGAAATTTGACCTCCAAAGGTTTAC 1416
Qy      401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      1417 CACCTGTCCAGCTGAGAAAGAGAGGCCATCCGGGTCTGTCTCTTTGATGGAATC 1476
Qy      421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      1477 GCTACAGGGCTCTGTGTGCTGAAGACTTGAGGCACTTCAAGTGAACCGCTACATTGCTCG 1536
Qy      441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db      1537 GAGGTGTGAGAGACTCCATCAACGGTGGGACATGTGCGCACAGGGGAAAGATCATGTAC 1596
Qy      461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
Db      1597 GTGGGGAGCTCCGACGCTCACAGAGCATATCAGAGAGTGGGCCCATTCGATCTG 1656
Qy      481 ValIleGlyIleSerProCysAsnAspLeuSerIleValAsnProAlaArgGlyGlyLeu 500
Db      1657 GTGATTTGGGGACATCCCTGCAATGACCTTCATGTCAACCTGCTCGCAAGGGCCTC 1716
Qy      501 TyrGluGlyThrGlyArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db      1717 TACGAGGGCAGTCGGCGGCTCTTCTTGTGATTCACCGGCTCTGATATGCGCGCCC 1776
Qy      521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
Db      1777 AAGAGAGAGATGATGCCCTTCTTGTGCTCTTGTGAGAAATGTGTGGCCATGGCGCTT 1836
Qy      541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db      1837 AGTGACAAGAGGAGCATCTCGGATTTCTCGAGTCCAAACCTGTGATGATGATGACCAA 1896
Qy      561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db      1897 GAAAGTCAAGCTGACACAGAGGCCCGCTAATTCTGGGGTAAACCTTCCCGTATGAACAG 1956
Qy      581 ProLeuAlaSerThrValAsnAspLysLeuGluGlnGluCysLeuGluHisIleGlyArg 600
Db      1957 CCGTTGCACTCACTGTGATATGATTAAGCTGAGAGCTGAGAGATGCTGAGACATGGCAG 2016
Qy      601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db      2017 ATAGCCAAGTTCAGCAAAAGTGAAGACATTACTACAGAGTCAAACTCCATAAGCAGGCG 2076
Qy      621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db      2077 AAAGACACGACATTTCTGTCTTCATGAATGAAAGAGACATCTTAAGTGTGACATGAA 2136
Qy      641 MetGluAspValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2137 ATGAAAGGATTTGTGTTTCCAGTCCACTTACTGAGAGTCTCAACATGAGCCCTTG 2196
Qy      661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db      2197 GCGAGGACAGAGACTGTGGGCGGTCAATGAGACGTGCCAGTCACTCGCACCTTTCGCT 2256
Qy      681 ProLeuLysGluTyrPheAlaCysVal 689
Db      2257 CCGTGAAGAGTATTTGCGGTGTG 2283

```

RESULT 6
 US-10-144-577-2
 ; Sequence 2, Application US/10144577
 ; Publication No. US20030083292A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MacLeod, Alan Robert
 ; TITLE OF INVENTION: Inhibitors of DNA Methyltransferase Isoforms

FILE REFERENCE: MET-005
 CURRENT APPLICATION NUMBER: US/10/144,577
 CURRENT FILING DATE: 2002-05-13
 PRIOR APPLICATION NUMBER: US 60/290,202
 PRIOR FILING DATE: 2001-05-11
 PRIOR APPLICATION NUMBER: US 60/290,212
 PRIOR FILING DATE: 2001-05-11
 NUMBER OF SEQ ID NOS: 49
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 2
 LENGTH: 3005
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-144-577-2

Alignment Scores:

Pred. No.:	0	Length:	3005
Score:	3714.00	Matches:	679
Percent Similarity:	98.8%	Conservative:	2
Best Local Similarity:	98.5%	Mismatches:	8
Query Match:	98.8%	Indels:	0
DB:	6	Gaps:	0

US-10-623-813-85 (1-689) X US-10-144-577-2 (1-3005)

QY 1 MetAsnAlaValGluGlnGlnAlaSerGlyGluSerGlnValGluGlnAlaSer 20
 DB 899 ATGAAATGCTGTGCAAGAAACCAAGGGCCCGGGAGCTCTCAAGAGTGGAGGGCCAGC 958
 QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 DB 959 CCGCTGCTGTGAGCAGCCAGCCAGCCGACATCCCGACCTGGGCTTACCAAGCTTGAG 1018
 QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 DB 1019 CCGGTGGGGTCCGATGCTGGGAGCAAGATGCCCAAGCAGCGCATGACGAGCCAGAG 1078
 QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
 DB 1079 TACGAGAGACCGCGCGGCTTTGGCATTTGGGAGCTGTGTGGGAAACTGGCGGCTTC 1138
 QY 81 SerTrpTrpProGlyArgGlyLeuSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
 DB 1139 TCTGTGGCCAGGCCCATTTGTCTTGTGATGACGGGCCGAGCCAGCAGCTGAA 1198
 QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
 DB 1199 GGCACCCGCTGGGTCATGTGGTTCGAGACGCGCAAAATTCACGTGTGTGAGAG 1258
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 DB 1259 CTATGCGCGTGAAGCTGCTTTTGCAGTGCCTTCCACAGGCCACAGTCAACAGCACCCC 1318
 QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 DB 1319 ATGTACCGCAAGCCATCTTACAGAGCTCTGCAAGTGGCCAGACCCCGCGGGAACTG 1378
 QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsn 180
 DB 1379 TTCGCGGTGTGCCAGCAGAGATGATGACACTGCCAAGGCGGTGAGGTGCAGAAC 1438
 QY 181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
 DB 1439 AACCCCTGATTTGATGATGGCCCTGGGGGCTTCCAGCTTCTGGCCCTTAAGGGCTTGA 1498
 QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
 DB 1499 CCACCAAGAGAGAGAAATCCCTACAAAGAGTGTACAGGACATCTGGGTGAACT 1558
 QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
 DB 1559 GAGGACACTGCTACGACCACTCCACAGCCAAAGAGCCCGGAGAGAGACAGCGGAG 1618

QY 241 LysProLysValIleGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
 DB 1619 AAGCCCAAGGTCAAGAGATTTATGTAGCGCACAGAGAGCGCTGTGTACAGGTG 1678
 QY 261 ArgGlnLysCysArgAsnIleGluIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 1679 CGGCAGAAATGCCGGAACATTGAGACATCTGATCTCTGTGGAGCTCAATGTTACC 1738
 QY 281 LeuGlnHisProLeuPheIleGlyIleMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
 DB 1739 CTGGAACACCCCTCTCTTGTGGAGAAATGTGCAAAATCTGCAGAAATGCTTTCGAG 1798
 QY 301 CysAlaTyrGlyThrTrpAspAspGlyTyrGlnIleTyrCysThrIleCysCysGly 320
 DB 1799 TGTGCTACCAAGTACAGACACAGCGGCTACAGCTCTTACCACTTGTCTGTGGGGC 1858
 QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysArgCysPheCysValGluCysVal 340
 DB 1859 CGTAGGTCTCATATGGCGGAAACAAACAACTGTCAGAGTCTTTGCGTGGAGTGTG 1918
 QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
 DB 1919 GACCTTGTGTGGGCGCGGGGCTGCGCCAGCGCATTAAGGAAACCCCTGGAACCTGC 1978
 QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGluAspTrpProSer 380
 DB 1979 TACATGTGGGGACAAAGGATCTACGGGCTCTCGCGGGAGAGACTGGCCCTCC 2038
 QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProLysValTyr 400
 DB 2039 CGGCTCAAGTGTCTTCGCTAATAACACAGACCAAGAAATTTACCTCCAAAGGTTAC 2098
 QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 DB 2099 CCACCTGTCCACCTTAGAAGAGAGCCCATCGGTGTGTCTCTTTGATGGAATC 2158
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
 DB 2159 GCTACAGGGCTCTGTGTCTGAAGACTTGGGCAATTCAGGTGACCGCTACATTTGCTCG 2218
 QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 DB 2219 GAGGTGTGAGACCTCATACAGGTGGCATGTGCGGACCAAGGAGATCATCTAC 2278
 QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
 DB 2279 GTGGGAGCTCCGACGCGTCAACAGAGCATATCCAGAGTGGGCCCATTCGATCTG 2338
 QY 481 ValIleGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2339 GTATTTGGGGAGTCCCTGTGCAATGACTTCCATGTCTCAACCTCTGCAAGGGCTC 2398
 QY 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
 DB 2399 TACGAGGACCTGGCGGCTCTTCTTGAGTTCTACCGCTCTGCAATGATGGCGGCC 2458
 QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaMetGlyVal 540
 DB 2459 AAGGAGAGAGATATGCCCTTCTTGTGCTTTGAGATGTGGTGGCCAGAGGGCGT 2518
 QY 541 SerAspLysArgAspIleSerArgPheLeuGluLysAsnProValIleAlaAspAlaLys 560
 DB 2519 AGTGACAGAGGACATCTGCGAATTTCTCGAGTCCAAACCTGTGATGATGACCAA 2578
 QY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArg 580
 DB 2579 GAGGTGACCTCAACACAGGCGCCGCTACTTCTGGGGTAACTTCCCGGTAAAGAACAG 2638
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGluGlnGluCysLeuGluHisGlyArg 600
 DB 2639 CCGTTGGCATCCACTGTGATGATTAAGCTGAGACTCAGAGAGTCTGGAGCATGGCAGG 2698
 QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620

Accession	Protein	Length	Score
Db	2699 ATAGCCAACTTCAGCAAAAGTAGAGCCATTACTACGAGTCCAAATCCATTAAGACAGGC	2758	
Qy	621 LysaSPG1NH1sPheProValPheMetLeuGluValysGluAspIleLeuTrpCysThrGlu	640	
Db	2759 AAAGCCACGACATTTCTCTCTCTTTCAGATGAGAAAGAGACATTTATGGTGCACTGA	2818	
Qy	641 MetGluArgValPheGluPheProValHisTyrThrAspValSerAsnMetSerArgLeu	660	
Db	2819 ATGGAAAGGGATATTTGGTTTCCCATCCACTATATCTGACGCTCCAAACATAGCCGCTTG	2878	
Qy	661 AlaArgGluArgLeuLeuGluValArgSerTrpSerValProValIleArgHisLeuPheAla	680	
Db	2879 GCGAGGCGAGAACACTCTGGCCCGGTCATGAGACCGTCCACATCATCCGACACTCTTCCCT	2938	
Qy	681 ProLeuLysGluTyrPheAlaCysVal	689	
Db	2939 CCGCTGAAGAGATATTTTGGCGGTGTGG	2965	

RESULT 7

```

US-10-172-118-132
? Sequence 132, Application US/10172118
? Publication No. US20030224374A1
? GENERAL INFORMATION:
? APPLICANT: Dai, Hongyue
? APPLICANT: He, Yudong
? APPLICANT: Linsley, Peter
? APPLICANT: Mao, Mao
? APPLICANT: Roberts, Chris
? APPLICANT: Van 't Veer, Laura
? APPLICANT: Van de Vijver, Marc
? APPLICANT: Bernards, Rene
? TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
? FILE REFERENCE: 9301-175-999
? CURRENT APPLICATION NUMBER: US/10/172,118
? CURRENT FILING DATE: 2002-06-14
? PRIOR APPLICATION NUMBER: 60/380, 770
? PRIOR FILING DATE: 2002-05-14
? NUMBER OF SEQ ID NOS: 2659
? SEQ ID NO 132
? LENGTH: 3005
? TYPE: DNA
? ORGANISM: Homo sapiens
? PUBLICATION INFORMATION:
? DATABASE ACCESSION NUMBER: AF067972
? DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-132

```

```

QY 441 GluValCysegluaspseriliethrvalglmetvalarghisglnglylmettyr 460
DB 2219 GAGGTGTGTAGGACCTTCATACAGGTGGCATGTGGGACCAAGGGAAGATCATGTAC 2278
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGlnLysPheProPheAspLeu 480
DB 2279 GTGGGGGACGTCCGACGCGTCACACAGAACATATCCAGAGAGTGGGGCCATTGCATCTG 2338
QY 481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
DB 2339 GTGATTGGGGGACGTCCCTGCATGACCTCTCCATGCTCAACCTGCTCCGACAGGCGCTC 2398
QY 501 TyrGlnGlyThrGlnLysArgLeuPheGlnPheThrArgLeuLeuHisAspAlaArgPro 520
DB 2399 TAGGAGGCACTGGCGGCTCTTCTTGAAGTTACCGCTCTCGCATGATGCGGAGCC 2458
QY 521 LysGlyGlyAspAspArgProPhePheTrpLeuPheGlnAsnValAlaIleMetGlyVal 540
DB 2459 AAGGAGGAGATATCCGCCCTTCTTGCTCTTGAGAAATGGTGGCCATGGGCGTT 2518
QY 541 SerAspLysArgAspIleSerArgPheLeuGlnSerAsnProValMetIleAspAlaLys 560
DB 2519 AGTGACAAGAGGACATCTCGCATTTCTCGAGTCCCAACCTGCTGATGATTGATGCCAA 2578
QY 561 GluValSerAlaIleHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetLysArg 580
DB 2579 GAAGTGTCAAGTGCACACAGGCGCGCTACTCTTGGGGTAACTCCGGATGAACAG 2638
QY 581 ProLeuAlaSerThrValAsnAspLysLeuGlnGlnLysCysLeuGlnHisIleGlyArg 600
DB 2639 CGCTTGCACTCCACTGTGATGATGATAGCTGAGCTCGAGAGTGTGAGCACTGCGAG 2698
QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
DB 2699 ATAGCCAAAGTTCAGCAAGAGGACCATTACTACGAGGTCAAACTCCATTAAGCGGGC 2758
QY 621 LysAspGlnHisPheProValPheMetAsnGlnLysGlnAspIleLeuTrpCysThrGln 640
DB 2759 AAGGACCGAGATTTCTGCTTCAATGAAAGAAAGAGACATCTTAATGGTCACTGAA 2818
QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
DB 2819 ATGGAAAGGGTATTTGGTTTCCAGTCCACTATACGAGCTCCCAACATGAGCCGCTTG 2878
QY 661 AlaArgGlnArgLeuGlnArgSerTrpSerValProValIleArgHisLeuPheAla 680
DB 2879 GCGAGGCGAGAGCATGCTGGGCGCGGATGAGAGCGTGCAGTCCGCCACTTCCGCT 2938
QY 681 ProLeuLysGlnTyrPheAlaCysVal 689
DB 2939 CCGCTGAAGAGATTTTGGCTGTGTG 2965

```

```

; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 132
; LENGTH: 3005
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-132

Alignment Scores:
Pred. No.: 0 Length: 3005
Score: 3714.00 Matches: 679
Percent Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatches: 8
Query Match: 98.8% Indels: 0
Gaps: 0

US-10-623-813-85 (1-689) x US-10-342-887-132 (1-3005)
QY 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyGlnSerGlnLysValGluGluAlaSer 20.
DB 899 ATGAATGCTGTGGAAGAAACCAAGGGGCCCGGGAGTCTCAAGAGTGGAGAGGCCAC 958
QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
DB 959 CCTCTGCTGTGCACACGCCCACTGACCCCGCATCCCACTGTGCTACACGCGCTGAG 1018
QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspLeuProGlu 60
DB 1019 CCGGTGGGGTCCGAGTGTGGGAGCAAGATGCAACAAAGCAGGCGATGACGAGCCAGAG 1078
QY 61 TyrGluAspGlyArgGlyPheGlyTyrIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
DB 1079 TAGAGAGAGGCGGGGGCTTGGCATGTGGAGCTGTGTGGGAAACTGCGGGCTTC 1138
QY 81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaIleGlu 100
DB 1139 TCCTGTGGCCAGGCGCGGATGTGCTTGAGAGTGAAGCGGCGGAGCGAGCACTGAA 1198
QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValAlaCysValGluLys 120
DB 1199 GGCACCGCGTGGTATGTGTGTCGAGACCGCAAAATCTCGTGTGTGTGTGAAG 1258
QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
DB 1259 CTGATGCGCTGAGCTGCTTTTGCAGTGCCTTCCACAGGCCACTACAAAGACGCC 1318
QY 141 MetTyrArgLysAlaIleTyrGlnValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
DB 1319 ATGTACCCGAAAGCATCATGAGGTCCTGACAGTGGCGAGAGCGCGGGGAAGCTG 1378
QY 161 PheProAlaCysHisAspSerAspGlnSerAspSerGlyLysAlaValGluValAsn 180
DB 1379 TTCCCGGTGTGCACAGACGATAGAGTGCACCTGCCAAGGCGGTGAAGGTGAAGAC 1438
QY 181 LysGlnMetIleGlnTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGln 200
DB 1439 AAGCCCATGATTGATATGGCCCTTGGGGGCTTCCAGCTTCTGCGCCCTAAGGCGCTGAG 1498
QY 201 ProProGluGlnGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
DB 1499 CCACCAAGAAAGAAAGATCCCTACAAAGAGTACACGACATGTGGGTGAACCT 1558
QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
DB 1559 GAGGAGGCTGCTTACGACCACTCCACAGCCAAAGAGCCCGGAAGGCAACCGGAG 1618
QY 241 LysProLysValLysGlnIleIleAspGluArgTyrArgGluArgLysValTyrGluVal 260
DB 1619 AAGCCCAAGGTCAAGAGATTTATGATGAGCCCAAGAGAGCGGCTGTGTACAGAGTG 1678
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280

```

```

Db      1679 CGGCAAGAGTCGGCAACATTGAGACATCTGCTCTGTGGAGACCTCAATTGTTACC 1738
Qy      281 LEUGLNIHIERProleuPheIIleGIyGlyMetCysGlnAsnCysIysAsnCysPheLeuGlu 300
Db      1739 CTGGAAACACCCCTCTCTTGGAGAAATGTCCAAACATCGCAAAATCGCTTTCTGGAG 1798
Qy      301 CyaaIATyGlnTyAspAspAspGIyTyArgIleSerTyCysThrIleCysCysGIyGly 320
Db      1799 TGTGCTTACCACTAGACACACGCGCTACCACTCTTACCACTCTGCTGTGGGGG 1858
Qy      321 ArgGluValIleuMetCysGIyAsnAsnAsnCysCysArgCysPheCysValGIyCysVal 340
Db      1859 CGTGAAGTCTCTATGTGGCAAAACAACCTGTCAGAGGCTTTGGCTGGAGTGTG 1918
Qy      341 AAPLeuLeuValGIyProGIyValIleGlnIleAlaIleIleGlyGluAspProTyrPheCys 360
Db      1919 GACCTCTTGTGGGGGGGGGGGGCTGCGCCAGGCGCATTTAAGGAACCCCTGGAACCTGC 1978
Qy      361 TyMetCysGIyHIleYsGIyThrTyArgIyLeuLeuArgArgGIyAspTyrProSer 380
Db      1979 TACATGTGGGGGCAAGGGTACTTACGGGCTGCTGCGGGGAGAGAGAGAGAGAGAGAGAG 2038
Qy      381 ArgLeuGlnMetPhePheAlaAsnAsnHIAspGIyGluPheAspProTyrValIy 400
Db      2039 CGGCTCCAGATGTTCTTCCCTAATAACAAGACCAAGAAATTTGACCTCCAAAGGTTTAC 2098
Qy      401 ProProValProAlaGIyAspArgIyProIleArgValIleuSerIleuPheAspGIyIle 420
Db      2099 CCACCTGTCCTCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2158
Qy      421 AlaThrGIyLeuLeuValIleuYsAspLeuGIyIleGlnIleValAspArgTyrIleAlaSer 440
Db      2159 GCTACAGGGCTCTGCTGCTGAAGAGCTTGGGCAATTCAGGTGAGACCTGCTGCTGCTG 2218
Qy      441 GlnValCysGIyAspSerIleThrValGIyMetValArgHIleGlnIyIySileMetYr 460
Db      2219 GAGGTGTGTGAGAGACTCCATCAAGGTGGGATGTGTGCGGACCAAGGAGAGATCATGTAC 2278
Qy      461 ValGIyAspValArgSerValThrGlnIySHIleGlnIyIleThrGIyProPheAspLeu 480
Db      2279 GTGGGGGAGCTCCGCGAGCTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2338
Qy      481 ValIleGIyIleSerProCysAsnAspLeuSerIleValAspProAlaArgIyGlyLeu 500
Db      2339 GTGATTTGGGGGAGAGTCTCTGCAATGACCTCTCATGCTCAACCTCTGCGCAAGGCTTC 2398
Qy      501 TyArgIyGlyThrGIyArgLeuPheGluPheTyArgIleuLeuHIAspAlaArgPro 520
Db      2399 TACGAGGGGACAGTGGCGGCTCTTCTTGAATGTTCTACCGCTCTCATGATGGCGGCC 2458
Qy      521 LysGIyGlyAspAspArgProPhePheTyrPheuPheGluAsnValAlaIleMetGIyVal 540
Db      2459 AAGGAGGGGAGATGATCGCCCTCTTCTGCTCTTGAATAATGTGTGCGCATGGCGGCTT 2518
Qy      541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaIy 560
Db      2519 AGTGACAAAGAGGACATCTCGCATTTCTCGAGTCCAACCTGTGTGATGATGACCAAA 2578
Qy      561 GlnValSerAlaAlaHIAspAlaArgIyTyPheTyrGIyAsnLeuProGIyMetAsnArg 580
Db      2579 GAAAGTCTACGCTGCACACAGGGGCGGCTACTTGTGGGTAACTTCCCGGTATGAACAG 2638
Qy      581 ProLeuAlaSerThrValIleAsnAspLysLeuGluGlnIyCysLeuGlnIySHIleArg 600
Db      2639 CCGTGTGGCATTCACCTGTGATGATTAAGCTGAGCTGCAGAGAGTCTTGAAGCATGGCAG 2698
Qy      601 IleAlaIyPheSerLysValArgThrIleThrTyArgSerAsnSerIleYsGIyGlnIy 620
Db      2699 ATAGCCAACTTACGCAAAAGTGAAGACCATTTACTACAGGTCAAATCCCAATAAAGCGGG 2758
Qy      621 LysAspGlnHIAspPheProValPheMetAsnGluYsGIyAspIleuTyrCysThrGlu 640
Db      2759 AAAAGACCAACATTTCTGTCTTCAATGAAAGAGAGACATCTTATGTGTCACTGAA 2818

```

```

Qy      641 MetGluArgValPheGIyPheProValHIleTyThrAspValSerAsnMetSerArgLeu 660
Db      2819 ATGAAAGGGATTTGTTGTTTCCCACTCACTTACTGACGTCTCAACATGAGCCCTTG 2878
Qy      661 AlaArgGlnArgLeuLeuGIyArgSerTyrSerValProValIleArgHIleuPheAla 680
Db      2879 CGGAGCGAGAGACTGCTGGGCGGTCATGAGACGTGCGCAGTCATCCGCACTCTTCCT 2938
Qy      681 ProLeuYsGIyTyrPheAlaCysVal 689
Db      2939 CCGTGAAGAGACTATTTGCGGTGTG 2965

RESULT 9
US-10-330-773-279
; Sequence 279, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 279
; LENGTH: 3009
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-773-279

Alignment Scores:
Pred. No.: 0
Score: 3714.00
Percent Similarity: 98.8%
Best Local Similarity: 98.5%
Query Match: 98.8%
Length: 3009
Matches: 679
Conservative: 2
Mismatch: 8
Indels: 0
Gaps: 0

US-10-623-813-85 (1-689) x US-10-330-773-279 (1-3009)
Qy      1 MetAsnAlaValAlaGlnIleAsnGlnAlaSerGIyGluSerGlnIyValGIyGluAlaSer 20
Db      903 ATGAAATCTCTGTGAAAGAAACCAAGGGGCGCGGGAGTCTCAAAAGTGAAGAGGCGCAC 962
Qy      21 ProProAlaValAlaGlnIleProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db      963 CCTCTGCTGTGCACAGCCCACTGACCCCGCATCCCACTGTGCTACCAAGCTGTGAG 1022
Qy      41 ProValGIyGlyAspAlaGIyAspLysAsnAlaThrLysValAlaAspAspGIyProGlu 60
Db      1023 CCGGTGGGTCCGAGTCTGGGGGACAAGATGCAACCAAGCGGATGACAGGCCAGAG 1082
Qy      61 TyArgIleAspGIyArgGIyPheGIyIleGIyGluLeuValTyrGIyLysLeuArgGIyPhe 80
Db      1083 TACGAGAGAGCGCGGGGCTTGTGCAATGGGAGCTGTGTGGGGGAAACTGCGGGGCTTC 1142
Qy      81 SerTyrTyrProGIyArgIleValSerTyrTyrMetThrGIyArgSerArgAlaIleGlu 100
Db      1143 TCTGTGTGGCCAGCGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1202
Qy      101 GlnThrArgTyrValMetTyrPheGIyAspGIyLysPheSerValValCysValGIyLys 120
Db      1203 GGCACCCCGTGGGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1262
Qy      121 LeuMetProLeuSerSerPheCysSerAlaPheHIleGlnAlaThrTyAsnLysGlnPro 140
Db      1263 CTGATGCCGCTGAGTCTGTTTGGAGTGTTCACAGGCGCACGTAACAAGCAGGCC 1322
Qy      141 MetTyArgLysAlaIleTyGIyValLeuGlnIleValAlaSerSerArgAlaGIyLysLeu 160
Db      1323 ATGTACCCCAAGGCATCTACAGAGTCTGTGAGGTGGCAGAGCGCGCGGGGAGAGCTG 1382

```

```

QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyValAlaGluValGlnAsn 180
Db 1383 TTCCCGGTGTGCGACGACGAGTGAAGTGAACCTCCAAAGGCCGTGGAGGTGCAGAAC 1442
QY 181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db 1443 AACCCCATGATTAATGAGCCCTGGGGGGCTTCAGGCTCTTGACCTTAAGGGCTTGGAG 1502
QY 201 ProProGluGluGluLysAsnProGlyLysGlyValAlaTrpThrAspMetTrpValGluPro 220
Db 1503 CCACCAAGAAAGAAAGAAATCCCTTACAAAGAGTGAACAAGGACATGTGGGTGAACCT 1562
QY 221 GluAlaAlaAlaTrpAlaProProProProAlaLysLysProLysSerThrTrpGlu 240
Db 1563 GAGGAGCTGCTTACGACCACTCCACCAAAAAGCCCGGAGAGACACAGCGGAG 1622
QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTrpGluVal 260
Db 1623 AAGCCCAAGGTCAAGAGATTATGATGAGCGCAAGAGAGGGCTGGTGTACAGAGTG 1682
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db 1683 CGGCAAGAGTGGCGGAAACATTGAGGACATCTGCATCTCTGGGAGCCTCAATGTACC 1742
QY 281 LeuGlnHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db 1743 CTGGAAACACCCCTCTCTGTGGAGAAATGTGCCAAAATGCAAGAACTCTTCTGGAG 1802
QY 301 CysAlaIleTrpGlnTrpAspAspAspGlyTrpGlnSerTrpCysThrIleCysCysGly 320
Db 1803 TGGCGGTACAGTACGACGAGAGAGGCTTACCACTCTCACTGCACTCTCTGTGGGGC 1862
QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db 1863 CGTGAAGTGTCTCATGTGCGGAAACAACACTGCTGCGAGTCTTTGCGTGAAGTGTG 1922
QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
Db 1923 GACCTCTGTGTGGGGCGGGGGCTGCCCAAGGCCATTAAAGAGAACCCCTTGAACTGC 1982
QY 361 TyrMetCysGlyHisLysGlyTrpTrpGlyLeuLeuArgArgArgLysAspTrpProSer 380
Db 1983 TACATGTGCGGACAAAGGTACTTACGGGCTGCTGGCGCGAGAGACTGGCCCTCC 2042
QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
Db 2043 CGGCTCAAGATGTTCTTCGCTAATTAACCAAGCAAGAAATTTGACCTCCAAAGTTTAC 2102
QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 2103 CCACCTGTCCCACTGAGAGAGAACCCCATCGGGTGTGCTCTCTTGTGATGATC 2162
QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAsnSer 440
Db 2163 GCTACAGGGCTCTGCTGCTGAAGGACTTGGGCACTTGAAGTGAACGCTACATTTGCTCG 2222
QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db 2223 GAGGTGTGTAGGACTCATCAGGTGGCATGTGTGGGACCAAGGGAGATCATGTAC 2282
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnIleLysProLysPheAspLeu 480
Db 2283 GTGGGGAGCTCGCAGCGTCAACAGAACATATCCAGAGTGGGGCCCATTCGATCTG 2342
QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db 2343 GTGATTTGGGGGAGTCCCTGCAATGACTCTCATCTCAACCTGCTCCAGAGGGCTC 2402
QY 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db 2403 TACGAGGGCACTGGCGCTCTTCTTGAAGTTTCAACGCTCCGATGATGCGCGGCC 2462

```

```

QY 521 LysGluGlyAspAspArgProPheTrpLeuPheGluAsnValAlaIleMetGlyVal 540
Db 2463 AAGGAGGGAGATGATGCCCCCTTCTTGTGCTCTTTGAAGATGTGTGGCATGGCCGT 2522
QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db 2523 AGTGAACAGAGGAGACTCTCGGATTTCTCGATTCAGATCCAACTCTGTGATGTGCCAAA 2582
QY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db 2583 GAAGTGTAGCTGCACACAGGGCCCGCTACTCTGGGGTAAACCTTCCGATGAACAGG 2642
QY 581 ProLeuAlaSerThrValAsnAspLysLeuGluGlnGlyCysLeuGlnHisGlyArg 600
Db 2643 CGTTGGCATCCACGTGATGATTAAGTGAAGCTGACAGCTGACAGAGTGTGAGCATGGCAGG 2702
QY 601 IleAlaLysPheSerLysValArgThrIleThrArgSerAsnSerIleLysGlnGly 620
Db 2703 ATAGCCAACTTCAAGCAATGAGAGACCATTAATGAGAGGTCAAACTCAATAAGCAGGCG 2762
QY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db 2763 AAAGACCAACATTTCTCTCTTCAATGAATGAAGAAAGAGACATTTATGTGCTGAA 2822
QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db 2823 ATGGAAGGGATTTGTGTTCCAGTCCACTATGAGAGTCTCCCAATGAGCCCTTG 2882
QY 661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db 2883 GCGAGGCAAGACTGTGGCGGCTGATGAGAGCGTCCAGTATCCGCCACCTTCTGCT 2942
QY 681 ProLeuLysGluTrpPheAlaCysVal 689
Db 2943 CCGCTGAAGAGTATTTGCGTGTGTG 2969

```

RESULT 10

US-09-720-086-3

Sequence 3, Application US/09720086

Publication No. US20060084053A1

GENERAL INFORMATION:

APPLICANT: Li, En

APPLICANT: Okano, Masaki

APPLICANT: Xie, Shaoping

TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,

FILE REFERENCE: 0609.4560002

CURRENT APPLICATION NUMBER: US/09/720, 086

PRIOR FILING DATE: 2000-12-20

PRIOR APPLICATION NUMBER: PCT/US99/14373

PRIOR FILING DATE: 1998-06-25

PRIOR APPLICATION NUMBER: 60/093, 993

PRIOR FILING DATE: 1998-07-24

NUMBER OF SEQ ID NOS: 82

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 4293

TYPE: DNA

ORGANISM: Homo sapiens

US-09-720-086-3

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length: 4293

Matches: 679

Conservative: 2

Mismatches: 8

Indels: 0

Gaps: 0

US-10-623-813-85 (1-689) x US-09-720-086-3 (1-4293)

1 MetAspAlaValGluGluGlnGlnAlaSerGlyGluSerGlnIleValGluGluAlaSer 20
 Db ATGAATGCTGTGGAAAGAAACAGAGGGCCCGGGAGTCTCAGAAAGTGGAGGAGCCACG 958
 Oy 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 Db 959 CCTCCGCTGTGACGACGCCACTGACCCCGCATCCCCCACTGTGCTACACAGCCCTGAG 1018
 Oy 41 ProValGlyGlyAspAlaGlyAspAlaThrIleValAlaAlaAspAspGluProGlu 60
 Db 1019 CCGGTGGGGTCCGATCTGTGGGACAAAGATGCCAACAGAGCCGATGACGAGCCAGAG 1078
 Oy 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyIleAspArgGlyPhe 80
 Db 1079 TACGAGGAGCGCCCGGGCTTTGGCATTTGGGAGCTGTGTGGGGGAAACTCGGGGCTTC 1138
 Oy 81 SerTrpTrpProGlyArgGlyIleValSerTrpTrpMetThrGlyAspSerArgAlaGlu 100
 Db 1139 TCCTGTGGCCAGGCGCGCATTTGTCTTGTGGATGACGGGCGCGAGCCGACGACTGAA 1198
 Oy 101 GlyThrArgTrpValMetTrpPheGlyAspGlyIleAspSerValValCysValGluIle 120
 Db 1199 GGCACCCCGTGGGTCTGTGTGGAGATGGCAAAATCTCAGTGGTGTGTGTGAAG 1258
 Oy 121 LeuMetProLeuSerSerPheCysSerAlaPheIleGlnAlaThrTyrAsnIleGlnPro 140
 Db 1259 CTGATGCCCTGAGTCCGCTTTTGGAGTGGCTTCCACAGGCCACGTACAAAGACGCC 1318
 Oy 141 MetTyrArgIleValAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyIle 160
 Db 1319 ATGTACCGCAAGCCATCTACAGAGTCTGCAAGTGGCAGAGCGCGCGGGAAGCTG 1378
 Oy 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyIleValAlaIleValGlnAsn 180
 Db 1379 TTCCTGGTGTGCACGACAGCATGAGTGAACATGCCAACAGCCCTGAGAGGTGCGAAG 1438
 Oy 181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProIleGlyLeuGlu 200
 Db 1439 AAGCCCATGATTTGAATGGGCGCTTGGGGGCTTCCAGCCCTTTCGGCCCTAAAGGCCCTG 1498
 Oy 201 ProProGluGluGluIleAspAspProTyrIleGlyValIleTyrThrAspMetTrpValGluPro 220
 Db 1499 CCACCAAGAAAGAGAGAAATCCCTACAAAGAGTGTACACGACATGTGGGTGGAACCT 1558
 Oy 221 GluAlaAlaIleTyrAlaProProProProAlaIleValIleValProArgIleSerThrArg 240
 Db 1559 GAGGAGCTGCTAGCACACCTCCACAGCCAAAGACCCCGGAGAGACACAGCGGAG 1618
 Oy 241 LysProIleValIleGlyIleIleAspGluArgThrArgGluValIleTyrGluVal 260
 Db 1619 AAGCCCAAGTCAAGAGATTTATTTGATGAGCCCAAGAGAGCGCTGCTGTACAGAGTG 1678
 Oy 261 ArgGlnIleCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 Db 1679 CGGCAAGAGTGCCTGCGAACTTGTAGGACATCTCATCTCTGTGGGAGCTCATATGTACC 1738
 Oy 281 LeuGluHisProLeuPheIleGlyIleGlyMetCysGlnAsnCysIleAspGlyLeuGlu 300
 Db 1739 CTGGAAACCCCTCTCTGTGAGGAATGTCCAAACCTGCAAGAACTGCTTCTGGAG 1798
 Oy 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
 Db 1799 TGTGGTACCACTACGACGACGACGCTACAGTCTTACGACCATCTGCTGTGGGGC 1858
 Oy 321 ArgGluValLeuMetCysGlyIleAspAspAsnCysCysArgCysPheCysValGluCysVal 340
 Db 1859 CGTGAAGTGTCTCATGTGCGGAAACAAACATCTGAGGTGCTTTTGTGCTGAGAGTGTG 1918
 Oy 341 AspLeuLeuValGlyProGlyValAlaGlnAlaAlaIleIleGluAspProThrTrpAsn 360
 Db 1919 GACCTCTGTGGTGGGGCGGGGCTGCCAGGACGCTTTAAAGAAAGACCCCTGGAACTGC 1978
 Oy 361 TyrMetCysGlyIleIleIleGlyIleTyrGlyLeuLeuArgArgGluAspTrpProSer 380

1979 TACATGTGGGCGCAAGGTACTTACGGGCTGTGGCGCGGCGAGAGGACTGGCCCTCC 2038
 Oy 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProIleValIle 400
 Db 2039 CGGCTTCCAGATTTCTTGGCTATATACCAAGACCAAGAAATTTGACCTTCAAGGTTTAC 2098
 Oy 401 ProProValProAlaGlyIleAspArgIleProIleArgValLeuSerIlePheAspGlyIle 420
 Db 2099 CCACTGTCTCCAGCTGAGAAAGAGAGCCCATCCGGGTCTGTCTCTTTGATGGAATC 2158
 Oy 421 AlaThrGlyLeuLeuValLeuIleAspLeuGlyIleGlnValAlaAspArgTyrIleAlaSer 440
 Db 2159 GCTACAGGGCTCTGTGTGCTGAAGACATTTGGGACTTGCAGGTGAGACCGCTACATTCCTCG 2218
 Oy 441 GluValCysGluAspSerIleThrValGluMetValArgHisGlnGlyIleIleMetTyr 460
 Db 2219 GAGGTGTGAGAGACTTCCATCACGGTGGCAGTGTGCGCACAGGGGAAAGATCATGTAC 2278
 Oy 461 ValGlyAspValArgSerValThrGlnIleHisIleGlnGluTrpGlyProPheAspLeu 480
 Db 2279 GTGGGGAGCTCCGACGCTCACAGAAAGCATATCCAGAGTGGGCCCCATTCGATCTG 2338
 Oy 481 ValIleGlyIleSerProCysAsnAspLeuSerIleValAsnProAlaArgIleGlyLeu 500
 Db 2339 GTGATGGGGGAGTCCCTGCAATGACTCTTCATCTCAACCTGCTGCAAGGGCCTC 2398
 Oy 501 TyrGluGlyThrGlyValArgLeuPheGluIlePheTyrArgGluLeuHisAspAlaArgPro 520
 Db 2399 TACGAGGCGACAGCGCGGCTCTTCTTGAATTTACCGCTCTCAGATATGCGGGCCC 2458
 Oy 521 LysGluGluAspAspArgProPhePheTrpLeuPheGluAsnValAlaAlaMetGlyVal 540
 Db 2459 AAGAGGAGATGATCGCCCTCTTCTGTGGCTTTTGAGAAATGTGTGGCCATTTGGCGCTT 2518
 Oy 541 SerAspIleArgAspIleSerIleArgPheLeuGluSerAspProValMetIleAspAlaIle 560
 Db 2519 AGTACAAAGAGGAGATCTCGGATTTCTCGAGTCCAAACCTGTGATGATGATCCAA 2578
 Oy 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAspArg 580
 Db 2579 GAAAGTACGTGACACACAGGCGCGCTTACTTGTGGGTAACTTCCGGTATGAACAGG 2638
 Oy 581 ProLeuAlaSerThrValAsnAspIleAspIleLeuGlnGluCysLeuGluHisIleGlyArg 600
 Db 2639 CCGTTGGCATCCATCTGTAATGATTAAGCTGAGAGTGTGAGAGTGTGAGAGATGGCAGG 2698
 Oy 601 IleAlaIlePheSerIleValArgThrIleThrThrArgSerAsnSerIleIleGlnGly 620
 Db 2699 ATAGCCAAAGTTCAGCAAGAGTGAAGACATTTACTACGAGGTCAAACCTCAATAAGCAGGCG 2758
 Oy 621 LysAspGlnHisPheProValPheMetAsnGluIleGluAspIleLeuTrpCysThrGlu 640
 Db 2759 AAAGACAGCATTTTCCGTCTTCAATGATGAGAAAGAGACATCTTATGTGGCTGA 2818
 Oy 641 MetGluAspValPheGlyPheProValHisTyrThrAspValSerAspMetSerArgLeu 660
 Db 2819 ATGAAAGGATTTTGGTTTCCAGTCCACTATATCTGACGTCTCAACATGAGCGGCTTG 2878
 Oy 661 AlaArgGlnArgLeuGluIleArgSerTrpSerValProValIleArgHisLeuPheAla 680
 Db 2879 GCGAGGACAGAGCTGTGGCGCGGCTCATGAGCGTGCAGCATATCGCCACCTTTCGCT 2938
 Oy 681 ProLeuIleGlyIlePheAlaCysVal 689
 Db 2939 CCGCTGAAGAGATTTTGGCGGTGTG 2965

RESULT 11

US-10-623-813-3

; Sequence 3, Application US/10623813

; Publication No. US20040234997A1

; GENERAL INFORMATION:

; APPLICANT: LI, En

```

APPLICANT: Okano, Masaki
APPLICANT: Xie, Shaoping
APPLICANT: Chen, Taiding
TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
TITLE OF INVENTION: Thereof
FILE REFERENCE: 0609,4580003
CURRENT APPLICATION NUMBER: US/10/623,813
CURRENT FILING DATE: 2003-07-22
PRIOR APPLICATION NUMBER: US 09/720,086
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: PCT/US99/14373
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 60/090,906
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: US 60/093,993
PRIOR FILING DATE: 1998-07-24
NUMBER OF SEQ ID NOS: 119
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 4293
TYPE: DNA
ORGANISM: Homo sapiens
US-10-623-813-3

Alignment Scores:
Pred. No.: 0 Length: 4293
Score: 3714.00 Matches: 679
Percent Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatches: 8
Query Match: 98.8% Indels: 0
DB: Gaps: 0

US-10-623-813-85 (1-689) x US-10-623-813-3 (1-4293)
QY 1 MetAenAlValGluGluValuSngInAlaserGlyGluSerGlnuValuGluGluAlaser 20
DB 899 ATGAATGCTGTGAGAAACCAAGGGCCCGGGAGTCTCAGAGGTGAGAGGGCCAGC 958
QY 21 ProProAlValGlnGlnProThraSPProAlaSerProThraValAlaThrThraProGlu 40
DB 959 CCTCCGCTGTGACAGCCCACTGACCCCGCATCCCACTGCTGCTACACAGCCCTGAG 1018
QY 41 ProValGlyGlyAspAlaGlyAspGlyAsnAlaThraValAlaAlaSPAspGluProGlu 60
DB 1019 CCCGTGGGGCTCCGATCTGGGGGCAAGAAATGCCAACCAAGCAGCATGACGAGCCAGAG 1078
QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
DB 1079 TACGAGGACGGCCGGGGCTTTGGCATTTGGGAGCTGGTGTGGGGAAACTGGGGGCTTC 1138
QY 81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaIleGlu 100
DB 1139 TCTGTGGGCGAGCCGCGCATTTGTCTGTGTGATGACGGGCGGAGCCGAGCAGCTGAA 1198
QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
DB 1199 GGCACCCCGCTGGGCTCATGTGGTTCCGAGACGCGCAAAATTTCTCAATGTGTGTGTGAGAG 1258
QY 121 LeuMetProLysSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
DB 1259 CTGATCCCGTGGCTGCTTTTTCAGTGTCTTCCACCAAGGCCACGTRCAACCAAGCAGCCC 1318
QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaIleLysLeu 160
DB 1319 ATGTACCGCAACCATCTACGAGGTCCTCAGAGTGGCCAGCAGCCGCGGGGAAAGCTG 1378
QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsn 180
DB 1379 TTCCCGGTGGCCACGACGATGAGAGTACACTGCCAAGGCCCGGTGAGGTGCAGAAC 1438
QY 181 LysGluMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
DB 1439 AAGCCCATGATTAATGGGCCCTTGGGGGCTTCCAGCCTTCTGAGCCTTAAGGGCCTGGAG 1498

```

```

QY 201 ProProGluGluGluLysAsnProTyrIleGlyValIleTyrThraSPMetTrpValGluPro 220
DB 1499 CCACCAAGAGAGAGAGAAATCCCTACAAAGAGGTACACGAGCATGTGGGTGAACTT 1558
QY 221 GluAlaAlaIleTyrAlaProProProProAlaLysLysProArgLysSerThrThraGlu 240
DB 1559 GAGGAGCTGCTTACGACCACTTCCACCGCAAAACCCCGGAAGGACACGCGGAG 1618
QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValIleTyrGluVal 260
DB 1619 AAGCCCAAGGTCAAGAGATTATTGATGAGCCCAAGAGAGACGGCTGTTACGAGGTG 1678
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
DB 1679 CGGCAAGAGTGGCAACATTAGACATCTGCATCTCTGTGGAGCTCTCAATGTTACC 1738
QY 281 LeuGluHisProLeuPheIleGlyLysMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
DB 1739 CTGAAACCCCTCTTCTTGTGAGGAATGTGCCAAACTGCAAGAACTTTCTTGAG 1798
QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
DB 1799 TGTGGGTACCACTACGACGACGAGTACCACTCTCACTGACCATCTGTGGGGGC 1858
QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
DB 1859 CGTGAGGTGCTCATGTGGGAAACAAACATGCTCAGAGTCTTTGCTGAGAGTGTG 1918
QY 341 AspLeuLeuValGlyProGlyValAlaGlnAlaIleLysGluAspProTrpAsnCys 360
DB 1919 GACCTCTGTGTGGGCGGGGGCTCCAGAGCATTAAAGAAAGACCCCTGGAATCTG 1978
QY 361 TyrMetCysGlyValLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
DB 1979 TACATGTGGGGGCAAGGGGTCTACGCGGTGCTGGCGCGGAGAGAGATCGCCCTCC 2038
QY 381 ArgLeuGlnMetPhePheIleAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
DB 2039 CGGCTTCAGATTTCTCTGCTAATACCAACGACGAAATTTGACCTCTCAAGTTTAC 2098
QY 401 ProProValProAlaGlyLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
DB 2099 CCACTGTCCAGCTGAGAAAGAAAGCCCATCCGGGTCTGTCTCTTGTATGGAATC 2158
QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAlaSPArgTyrIleAlaSer 440
DB 2159 GCTACAGGGCTCTGCTGTGAAGACTTGGGCATTCAAGGTGAGCCGCTACATTGCTCG 2218
QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
DB 2219 GAGGTGTGAGAGACTCTCATCCAGGTGGGCAAGTGGGACCAAGGAAATCATGTAC 2278
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
DB 2279 GTGGGGAGCTCCGACGCTCACACAGAAAGCATATCCAGAGTGGGGCCATTGATCTG 2338
QY 481 ValIleGlyIleSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
DB 2339 GTGATGGGGGACGTCCTGCAATGACCTTCATCTGTCACACCTGCTGCAAGGGCTTC 2398
QY 501 TyrGluGlyThrGlyArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
DB 2399 TACGAGGGACAGCGCGGCTCTTCTTGAGTTCTACCGCTCTCTGATATGCGCGGCC 2458
QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
DB 2459 AAGAGAGGAGATGATGCGCCCTTCTTGTGCTTGTGAAGATGTGTGACCATGGCGCTT 2518
QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
DB 2519 AGTGACAAAGAGGACATCTCGGATTTCTGAGATCCCAACCTGTGTATGATTGTCAAA 2578

```

QY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGluValAsnLeuProGluMetAsnArg 580
 DB 2579 GAAGTGTACAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTAAGACAGG 2638
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
 DB 2639 CCGTTGGCATCCACTGTGAATGATTAAGCTGGAGCTGCAGAGGTGTCTGGAGCATGGCAGG 2698
 QY 601 ILAAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
 DB 2699 ATGGCCAAAGTTCAGCAAGAGAGACCATTAACGAGTCAACCTCCATAAAGCAGGGC 2758
 QY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTyrCysThrGlu 640
 DB 2759 AAAGACCAAGATTTTCTGTCTTCATGAATGAGAAAGAGACATCTTATGTGCATCGAA 2818
 QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 DB 2819 ATGAAAGGATATTGGTTCCAGTCCACTATACCTTCACATGAGCGCGCTTG 2878
 QY 661 AlaArgGlnArgLeuGluArgSerTrpSerValProValIleArgHisLeuPheAla 680
 DB 2879 GCGAGGCAAGAGACTGTGGCCGCTGCATGAGCGTGCAGTCACTCCGCCACTTTCGCT 2938
 QY 681 ProLeuLysGluTyrPheAlaCysVal 689
 DB 2939 CCGCTGAAGGAGTATTTCGCGTGTG 2965

RESULT 12
 US-10-264-237-953
 / Sequence 953, Application US/10264237
 / Publication No. US20040009491A1
 / GENERAL INFORMATION:
 / APPLICANT: Birst et al.
 / TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 / FILE REFERENCE: PA131P1
 / CURRENT APPLICATION NUMBER: US/10/264,237
 / CURRENT FILING DATE: 2002-10-04
 / PRIOR APPLICATION NUMBER: PCT/US01/16450
 / PRIOR FILING DATE: 2001-05-18
 / PRIOR APPLICATION NUMBER: US 60/205,515
 / PRIOR FILING DATE: 2000-05-19
 / NUMBER OF SEQ ID NOS: 2876
 / SOFTWARE: Patent In Ver. 3.1
 / SEQ ID NO 953
 / LENGTH: 2938
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: misc_feature
 / LOCATION: (491)..(491)
 / OTHER INFORMATION: n equals a,c,g, or c
 / FEATURE:
 / NAME/KEY: misc_feature
 / LOCATION: (1987)..(1987)
 / OTHER INFORMATION: n equals a,c,g, or c
 / FEATURE:
 / NAME/KEY: misc_feature
 / LOCATION: (1994)..(1994)
 / OTHER INFORMATION: n equals a,c,g, or c
 / FEATURE:
 / NAME/KEY: misc_feature
 / LOCATION: (2892)..(2892)
 / OTHER INFORMATION: n equals a,c,g, or c
 / FEATURE:
 / NAME/KEY: misc_feature
 / LOCATION: (2932)..(2932)
 / OTHER INFORMATION: n equals a,c,g, or c
 / US-10-264-237-953

Alignment Scores: 1.37e-297 Length: 2938
 Pred. No.: 2708.00 Matches: 494
 Score:

Percent Similarity: 98.8% Conservative: 1
 Best Local Similarity: 98.6% Mismatches: 6
 Query Match: 72.1% Indels: 0
 DB: 7 Gaps: 0
 US-10-623-813-85 (1-689) x US-10-264-237-953 (1-2938)
 QY 189 GLYGlyPheGlnProSerGlyProLysGlyLeuGluProProGluGluGlyLysAsnPro 208
 DB 1 GGCACAGCCAGCCTTCTGACCTTAAGGCTTGAAGCCACCAAGAGAGAAATCCC 60
 QY 209 TYLYSGluValTYrThrAspMetTrpValGluProGluAlaAlaIleTyrAlaProPro 228
 DB 61 TACAAAGAAAGTGAACGAGACATGTGGTGAACCTGAGGACAGCTGCTACGACACCT 120
 QY 229 ProProAlaLysLysProArgLysSerThrThrGluLysProLysValLysGluIleIle 248
 DB 121 CCACCAAGCCAAAAGCCCGGAGAGACACGCGGAAAGCCCAAGGTCAAGAGATTATT 180
 QY 249 AspGluArgThrArgGluArgLeuValTYrGluValArgGluLysCysArgAsnIleGlu 268
 DB 181 GATGAGCGACAAAGAGAGCGGCTGTGTACAGGTCCGCGAAGTCCGGAACATTGAG 240
 QY 269 AspIleCysIleSerCysGlySerLeuAsnValThrLeuGluHisProLeuPheIleGly 288
 DB 241 GACATCTGATCTCCCTGTGGAGCTCAATGTTCACCTGGAACACCCCTTCGTTGGA 300
 QY 289 GlyMetCysGlnAsnCysLysAsnCysPheLeuGluCysAlaTyrGlnTyrAspAspAsp 308
 DB 301 GGAATGTGCCAAAACGCAAGAACTGCTTCTGAGTGTGCGTACCAAGTACGACGAC 360
 QY 309 GLYTYrGlnSerTYrCysThrIleCysCysGlyGlyArgGluValLeuMetCysGlyAsn 328
 DB 361 GGCTACCACTCTTACGACCACTGTCTGTGGGGCCGTGAGTGTCTCATGTGCGAAAC 420
 QY 329 AsnAsnCysCysArgCysPheCysValGluCysValAspLeuLeuValGlyProGlyAla 348
 DB 421 AACCACTGTGACAGTGTCTTTCGCGAGTGTGTGACTCTTGTGGGGCCGGGGCT 480
 QY 349 AlaGlnAlaAlaIleLysGluAspProTrpAsnCysTYrMetCysGlyHisGlyGlyThr 368
 DB 481 GSCCARGCAGNACTTAAAGAAAGACCCCTGGAACCTGACATGTGCGGACCAAGGGTACC 540
 QY 369 TYrGlyLeuLeuArgArgGluAspTrpProSerArgLeuGlnMetPhePheAlaAsn 388
 DB 541 TACGGGCTCTCGGGGCGAGAGAGACTGCCCTCCGGTCCAGATGTTCTTCCTTAAT 600
 QY 389 AsnHisAspGlnGluPheAspProProLysValTYrProProValProAlaGluLysArg 408
 DB 601 AACCAAGACACAGAAATTGACCTCCAAAGGTTTACCACCTGTCCACGTGAGAAAGAG 660
 QY 409 LysProIleArgValLeuSerLeuPheAspGlyIleAlaThrGlyLeuLeuValLeuLys 428
 DB 661 AAGCCCATCCGGGTGTCTCTCTTTGATGGAATGCTTACAGGGCTCTCTGTGCTGAAG 720
 QY 429 AspLeuGlyIleGluValAspArgTYrIleAlaSerGluValCysGluAspSerIleThr 448
 DB 721 GACTTGGGCAATTCAGGTGACCGCTTACATTGCTCGAAGTGTGTAGAGATTCATCAGC 780
 QY 449 ValGlyMetValArgHisGlnGlyLysIleMetTYrValGlyAspValArgSerValThr 468
 DB 781 GTGGGATGTGTGGGACAGGGGGAAGATCATGTCGTGGGAGCTCCGACGCTCACCA 840
 QY 469 GlnLysHisIleGlnGluTyrPheLysAspLeuValIleGlyGlySerProCysAsn 488
 DB 841 CAGAAGCATATTCAGAGAGTGGGCCCATTCGATCTGTGATTGGGGGAGTCCCTGCAAT 900
 QY 489 AspLeuSerIleValAsnProAlaArgLysGlyLeuTYrGluGlyThrGlyArgLeuPhe 508
 DB 901 GACCTTCCATGTCATCAACCTCTCTGCAAGGGCTTACGAGGGCACTGGCGGCTCTTC 960
 QY 509 PheGluPheTYrArgLeuLeuHisAspAlaArgProLysGluGlyAspAspArgProPhe 528

Db	961	TTTGAGTTCTACCGCGCTCTCGATGATGCGGGGCCCAAGGAGGGAGATGATCGCCCTTC	1028
Qy	529	PheTrIePuePheGluAenValValAlaMeGlyValSerAspLYsArgAspLLeSerArg	548
Db	1021	TTCTGGCTCTTTGGAAATGTGTGGCCCTGAGCGCTTACTGTGCAAGAGGACATCTTCGGCA	1080
Qy	549	PheLeuGluSerAmPProValMetCLeAspAlaLYsGluValSerAlaAlaHisArgAla	568
Db	1081	TTTCTCGAGTCCAACTCTGTATGTATGTATGCTCAAGAGAGTCTACGTGCACACAGGGCC	1140
Qy	569	ArgTrYPheTrPGLYAsnLeuProGlyMetAsnArgProLeuAlaSerThrValAsnAsp	588
Db	1141	CGCTACTTCTTGGGTAACTCTCCCGTATGAAACAGGCCGTTGGCATTCACCTGTGAATGAT	1200
Qy	589	LYsLeuGluLeuGingluCYsLeuGluHisGlyArgIleAlaLYsPheSerLYsValArg	608
Db	1201	AAGCTGAGCGTCCAGGAGTGTCTGAGCATGCGACGATAGCCAAAGTTCCAGCAAGTAGAG	1260
Qy	609	ThrIleThrTrArgSerAsnSerLLeLYsGingLYsAspGlnHisPheProValPhe	628
Db	1261	ACCATTTACTACGAGAGTCAAACTCCATAAAGCAGGCAAGCAGCATTTTCTCTCTTC	1320
Qy	629	MetAsnGluLYsGluAspLLeuTrpCysThrGluMetGluuArgValPheGlyPhePro	648
Db	1321	ATGATAGTGAAGAAAGGAGACATTTATGTGTGCACCTGAAATGAGAAAGGATATTGGATTCCCA	1380
Qy	649	ValHisTrYPheTrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeuLeuGlyArg	668
Db	1381	GTCACACTACTGACGTCCTCCAACTAGAGCCGCTTGGGAGGACGAGACTCTCTGGCCGG	1440
Qy	669	SerTrpSerValProValIleArgHisLeuPheAlaProLeuLYsGluTrYPheAlaCys	688
Db	1441	TCATGAGAGCGTGCCAGTCATCCGCCACTTTCGCTCGCTGAAGAGATATTTCGCGTGT	1500
Qy	689	ValI 689	
Db	1501	GTG 1503	
RESULT 13			
US-09-720-086-4			
; Sequence 4, Application US/09720086			
; Publication No. US20060084053A1			
; GENERAL INFORMATION:			
; APPLICANT: Li, En			
; APPLICANT: Okano, Masaki			
; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,			
; FILE REFERENCE: 0609.456002			
; CURRENT APPLICATION NUMBER: US/09/720,086			
; PCT FILING DATE: 2000-12-20			
; PRIOR APPLICATION NUMBER: PCT/US99/14373			
; PRIOR FILING DATE: 1999-06-25			
; PRIOR APPLICATION NUMBER: 60/090,906			
; PRIOR FILING DATE: 1998-06-25			
; PRIOR APPLICATION NUMBER: 60/093,993			
; PRIOR FILING DATE: 1998-07-24			
; NUMBER OF SEQ ID NOS: 82			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 4			
; LENGTH: 4145			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; US-09-720-086-4			

Alignment Scores:	
Pred. No.:	1,09e-234
Score:	2159.50
Percent Similarity:	71.8%
Best Local Similarity:	56.5%
Query Match:	57.5%
DB:	5
Gaps:	.
	7
Length:	4145
Matches:	410
Conservative:	93
Mismatches:	158
Indels:	41

US-10-623-813-85 (1..689) x US-09-720-086-4 (1..4145)

Qy	10	SerGlyGlnSerGlnValGluGlnAlaSerProProAlaValGlnGlnProThrAsp	29
Db	623	TCACAGACGACACAGAGGACACACATGGACGCCCCAGACGACGATACCCCTTACG---	679
Qy	30	ProAlaSerProThrValAlaThr-----ThrProGlnProValGlyAsp	45
Db	680	CCCGCCTTACGCCAGACAGCCAGACGAGGGGACATGGAGTCCCGCAG-GTGAGAGCAGAC	738
Qy	46	AlaGlyAspIysAsnAlaThrIysAlaAlaAspAspGlnProGlnIYrGlnAspGlyArg	65
Db	739	AGTGGAGAT-----GGAGACGTTTCAGAGTATCAGAGTACGAGAGGGAAG	777
Qy	66	GlyPheGlyIleGlyGlnLeuValTTPGlyLysLeuArgGlyPheSerTTPProGly	85
Db	778	GAGTTTGGAAATAGGGACCTCGTGTGGGAAAGATCAAGGGCTTCTCTGTGGCCCCG	837
Qy	86	ArgIleValSerTTPTrpMetThrGlyArgSerArgAlaIleGlnGlyThrArgTTPVal	105
Db	838	ATGGTGTGTCTTGGAAAGCCACCTCCACGACGACCTATGTCTGGCATGCGGTGGTGC	897
Qy	106	MetTrpPheGlyAspGlyLysPheSerValValCysValGlnIYrLysLeuMetProLeuSer	125
Db	898	CAGGTGTGGCCATGGCAGAGTTCTCCAGGCTCTTCCAGACAAACTGTGTGGCATGGGG	957
Qy	126	SerPheCysSerAlaPheIleGlnAlaThrThrAsnLysGlnProMetTyrArgLysAla	145
Db	958	CTGTTCAGCAGACACTTATATTGGCCACCTTCATATAGCTCGCTCTTCGAAAGGCC	1017
Qy	146	IleTyrGlnValLeuGlnValAlaIleSerSerArgAlaGlyLysLeuPheProAlaCysHis	165
Db	1018	ATGATCCATGCTCTGGAGAAAGCTAGGGTCCGAGCTGGCAAGACTTCCCC-----	1068
Qy	166	AspSerAspIuSerAspSerGlyLysAlaValGlnValGlnAsnLysGlnMetIleGlu	185
Db	1069	-----AGCAGCCCTGGAGACTATTGGAGAGCAGCTGAAGCCCATGTTGGAG	1116
Qy	186	TTPAlaLeuGlyLysPheGlnProSerGlyProlGlyLeuGlnProProGlnGlu---	204
Db	1117	TGGGCCACCGGGGCTTCAAGCCCATCGGATGAGGGCTTCAAACTCAACACAGCCAA	1176
Qy	205	-----GlnLysAsnPro	208
Db	1177	CCAGTGGTAAATAAGTCGAAAGTCGCTCGTCGACGACGATGAGAAATAGAAATCAAGAA	1236
Qy	209	TyrLysGlnValTyrThrAspMetTTPValGlnProGlnAlaAla---TyrAlaPro	227
Db	1237	TACGAGAAACAAGACTCGAAGACCCACAGCTGACGACTCAGCCACTCTGACTGTGCC	1296
Qy	228	ProProProAlaLysIysProArgLysSerThrThrGlnLysProlLysValLysGluIle	247
Db	1297	GCAACCCAAAGCCCTCAAGACAAATTGCTATPAACAACGGCAAAAGACGAGGGGATGAA--	1353
Qy	248	IleAspGlnArgThrArgGlnArgLeuValTyrGlnValAlaArgGlnLysCysArgAsnIle	267
Db	1354	-----GATCAGAGCGGAGAAACAATGGCTTCAGATGTTGCCAAACAAGACAGAGCTG	1407
Qy	268	GlnAspIleCysIleSerCysGlySerLeuAsnValThrLeuGlnHisProlLeuPheIle	287
Db	1408	GAGATGCTGTGTGTCTTGTGGCAGGAAAAACCCCGTCTTCACCCCTCTTTGAG	1467
Qy	288	GlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlnCysAlaTyrGlnTyrAspAsp	307
Db	1468	GGGGGCTCTGTCTCAGACATCCCGGATCGCTTCTTACGCTGTTTACATATAGATGAC	1527
Qy	308	AspGlyTyrGlnSerTyrCysThrIleCysCysGlyLysArgGlnValLeuMetCysGly	327
Db	1528	GATGCTATCAGTCTTACTCACTGTGTCTGCGAGGCGCAGAGCTGCTGCTTTGACG	1587
Qy	328	AsnAsnAsnAsnCysCysArgCysPheCysValGlnCysValAspLeuLeuValGlyProGly	347
Db	1588	AATCAGGAGCTGCTCCGCTTTCTGTGTGTGAAGCTCTGAGAGTCTGTGTGGCAGAGC	1647

[illegible]

```

348 AlaAlaGlnAlaAlaIleLeuSgluAspProTyrAsnCysTyrMetCysGlyHisIleSglu 367
Db ACAGGGCCGAGAGCCAAAGCTTACGAGCCCTGGAGCTGCTACATGTCCTCCCGAGCC 1707
Qy ThrTyrGlyLeuLeuArgArgArgIleAspTyrProSerArgLeuGlnMetPhePheAla 387
Db TGTCAATGGCTCTGGCGCGCGGAAAGCATGTAAGCTGGCTGAGCCCTTCTTCAAC 1767
Qy AsnaAsnHisAspGlnGluPheAspProPolysValTyrProProValProAlaGluLys 407
Db AGTGACACGGCGCTTGAATACCAAGCCCCCAAGCTGACCTGACCTATCCCGAGCCGA 1827
Qy ArgLysProIleArgValLeuSerLeuPheAspGlyIleAlaIleThyGlyLeuValLeu 427
Db AGCGCGCCCATTCGAGTCTGATCTTGTGATGGCATGCAACGCTACCTAGTCTTC 1887
Qy LysAspLeuGlyIleGlnValAspArgTyrIleAlaSerGlyValCysGluAspSerIle 447
Db AAAGAGTTGGGATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1947
Qy ThrValGlyMetValArgHisGlnGlyLysIleMetTyrValGlyAspValArgSerVal 467
Db GCTGTTGGAACCGTGAAGCAGCAGGGAATATCAATACGTGAACGAGCTGAGAAATC 2007
Qy ThrGlnLysHisIleGlnGlnTyrProPheAspLeuValIleGlyGlySerProCys 487
Db ACMAAGAAAAATATGGAATAGGAGCCCATTTGATGCTGATGAGCGGAAAGCCCATGC 2067
Qy AsnaAspLeuSerIleValAsnProAlaArgLysGlyLeuTyrGlyIleThyGlyArgLeu 507
Db AACGATCTCTCAAAATGTGAATCCAGCCAGAAAGGCTGTATGAGGTCACAGCCGCTC 2127
Qy PhePheGluPheTyrArgLeuLeuHisAspAlaArgProLysGlnGlyAspAspArgPro 527
Db TTCTTCGAATTTTACACCTGCTGAATTACTCAAGCCCAAGAGGATGATACCGGCGG 2187
Qy PhePheTyrPhePheGlnGlnValValAlaMetGlyValSerAspLysArgAspIleSer 547
Db TTCTTCTGATGTTGAGATGTTGAGCCTGAAGAGTTGGCGCAAGGAGGACATCTTA 2247
Qy ArgPheLeuGluSerAsnProValMetIleAspAlaLysGluValSerAlaIleHisArg 567
Db CGGTTCTCGAGTGAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2307
Qy AlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArgProLeuAlaSerThrValAsn 587
Db GCCGATACCTCTGGGGCAACTACCCGGGATGAACAGGCCCGGTGATGATCAAAAGAT 2367
Qy AspLysLeuGluLeuGlnGlnCysLeuGlnHisGlyArgIleAlaLysPheSerVal 607
Db GATAAACTCGAGCTCAGAGCTGCTTGAATACATAGAGTACCAAGTAAAGAAAGTA 2427
Qy ArgThrIleThrThrArgSerAsnSerIleLysGlnGlyLysAspGlnHisPheProVal 627
Db CAGACAAATACCAACCAAGCTGAGTCACTCAACAGGGGAAACCAACTTTTCCCTGTT 2487
Qy PheMetAsnGluLysGluAspIleLeuTyrPyrThrGluMetGluArgValPheGlyPhe 647
Db GTCATGATGGCAAGAAAGATGTTTGTGTCACAGAGCTGAAAGGATCTTTGGCTTT 2547
Qy ProValHisTyrThrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeuGly 667
Db CCTGTGCACTACAGACGTCGTCAACATGGGCGGTGTCGCCGCAAGAGCTGCTGGGA 2607
Qy ArgSerTyrPseValProValIleArgHisLeuPheAlaProLeuLysGlnTyrPheAla 687
Db AGGTCTTGAGGCTGCTCTCTCATCCGACCTCTTCCCTCTGAGAGCACTATTGCA 2667
Qy Cys 688
Db TGT 2670

```

```

RESULT 14
US-10-144-577-3
; Sequence 3, Application US/10144577
; Publication No. US20030083292A1
; GENERAL INFORMATION:
; APPLICANT: MacLeod, Alan Robert
; TITLE OF INVENTION: Inhibitors of DNA Methyltransferase Isoforms
; FILE REFERENCE: MET-005
; CURRENT APPLICATION NUMBER: US/10/144,577
; PRIOR APPLICATION NUMBER: US 60/290,202
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,212
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4145
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-144-577-3

Alignment Scores:
Pred. No.: 1,096-234 Length: 4145
Score: 2159.50 Matches: 410
Percent Similarity: 71.84 Conservative: 93
Best Local Similarity: 58.58 Mismatches: 158
Query Match: 57.58 Indels: 41
Gaps: 7

US-10-623-813-85 (1-689) x US-10-144-577-3 (1-4145)
Qy 10 SerGlyGluSerGlnLysValGlnGluLysAspProProAlaValGlnGlnProThrAsp 29
Db 623 TCACAGCAGCAGACAGAGCAGACACATGGAGCCGCCAGACAGCAGTACCCCTACG--- 679
Qy 30 ProAlaSerProThrValAlaThr-----ThrProGluProValGlyLysAsp 45
Db 680 CCCGCTAGCCAGCAGCAGCAGCAGCAGGCGCATGATCCCGCAG--GTGAGCAGAGC 738
Qy 46 AlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGluTyrGlyLysGlyLysArg 65
Db 739 AGTGAGAGAT-----GGAGCAGCTTCAAGATATCAGATGATGAGTGGAG 777
Qy 66 GlyPheGlyIleGlnGlnLeuValTyrGlyLysLeuArgGlyPheSerTyrThrProGly 85
Db 778 GAGTTGGAATPAGGGGAGCTGTGTGGGAAAGATCAAGGCTTCTCTGGTGGCCGCC 837
Qy 86 ArgIleValSerTyrTyrPheMetThrLysArgSerArgAlaAlaGlnGlyThrArgTyrVal 105
Db 838 ATGTGTGTCTTGGAAAGCCACCTCCCAAGCAGCAGCATGTCTGGCATGGCGTGGGTC 897
Qy 106 MetTyrPheGlyAspGlyLysPheSerValCysValGluLysLeuMetProLeuSer 125
Db 898 CAGTGTGTGGCAGATGGCAAGTTCGAGGTCTCGACGACAAACTGTGTGCACCTGGGG 957
Qy 126 SerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnProMetTyrArgLysAla 145
Db 958 CTGTTACCCAGCAGCACTTTAATTTGGCCACTTCAATTAAGCTGTGTCTCCATGAAAGGC 1017
Qy 146 IleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeuPheProAlaCysHis 165
Db 1018 ATGTTCATGTCTTGGAAGAAAGTAAAGTGGAGTGGCAAGTGGCAAGCTTCCG--- 1068
Qy 166 AspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsnLysGlnMetIleGlu 185
Db 1069 -----AGCAGCCCTGGAGACTATGGAGGACCAAGCTGAAGCCCATGTTGGAG 1116
Qy 186 ThrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGluProProGluGlu--- 204
Db 1117 TGGGCCACGGGGGCTTCAAGCTCAAGCTGAGATTCAGAGGCTTCAACCAACACACCGCA 1176
Qy 205 -----GluLysAsnPro 208

```

Db	1177	CCAGTGGTTAACTCAAGAGTGGCTGGTGCAGGCACTGAGAAATTAGATTAACAAGAA	1236
Qy	209	TYTYSGLVAlaTYTThrAspMetTrpValGluProGluAlaAlaAla---TYTAlaPro	227
Db	1237	TACGGAAACAAGACTCGAAGACGCACAGCTGACGACTGACGACACTTCGACTACGCCCC	1296
Qy	228	ProPProAlaLysLysProArgLysSerThrTrpGluLysProLysValLysGluLe	247
Db	1297	GCACCAAGCCCTCAAGACAAATTGCTATACACGGCAAAAGACCGAGGGGATGAA---	1353
Qy	248	ILaspGluArgThrArgGluArgLeuValTyrGluValArgGlnLysCysArgAsnIle	267
Db	1354	-----GATCAGAGCCCGAAGAACAAATGCTTCAGATGTGGCAACAAAGAGACCTTG	1407
Qy	268	GluAspIleCysIleSerCysGlySerLeuAsnValThrLeuGlnHisProLeuPheIle	287
Db	1408	GAAAGTGGCTGTGTCTTGTGGCGAGAAAAACCCCGTCTCTTCACCCCTCTTTGAG	1467
Qy	288	GlyIleMetCysGlnAsnCysLysAsnCysPheLeuGluCysAlaTyrGlnTyrAspAsp	307
Db	1468	GGGGGGCTCTGTACAGACATGCCGGGATCGCTTCCTTGAGCTGTATTACATGATGATAC	1527
Qy	308	AspGlyTyrGlnSerTyrCysThrIleCysCysGlyArgGluValLeuMetCysGly	327
Db	1528	GATGGCTATACGCTTACTGTGACGTGTGCTGCGAGGGCCGAGAGCTGCTCTTGACGC	1587
Qy	328	AsnAsnAsnCysCysArgCysPheCysValGluCysValAspLeuValGluProGly	347
Db	1588	AACACAGACTCTCTCCGGTGTTCGTGTGAGAGTCCCTGGAGAGTCTGTGTGGCACAGCC	1647
Qy	348	AlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCysTyrMetCysGlyHisLysGly	367
Db	1648	ACACGGCGCGGAGCCAGACTTCAGAGAGCCCTGGAGCTGCTCAATGTCTCCGACAGGC	1707
Qy	368	ThrTyrGlyLeuLeuArgArgArgGluAspTrpProSerArgLysGlnMetPhePheAla	387
Db	1708	TGTCATGGCGTCTCTCGCGCGCCGCGAAGACCTGGAAACGTGCCCTTCGACGCTTTACCC	1767
Qy	388	AsnAsnHisAspGlnLysPheAspProProLysValTyrProProValProAlaGluLys	407
Db	1768	AGTACACAGGGGCTTGAATACGAAGCCCCCAAGCGTACCTCCGACTTCCGCGACCGGA	1827
Qy	408	ArgLysProIleArgValLeuSerLeuPheAspGlyIleAlaThrGlyLeuLeuValLeu	427
Db	1828	AGCGGGCCCATTCGAGTCTGTCAATGTTTATGACCATCGGACAGGCTTACTTATGCTCTC	1887
Qy	428	LysAspLeuGlyIleGlnValAspArgTyrIleAspSerGluValCysGluAspSerIle	447
Db	1888	AAAGAGTTGGGCATTAAGTAGTAAGTAAGTAAGTACGTCCTCTGAAGTGTGAAGATTCATT	1947
Qy	448	ThrValGluMetValArgHisGlnGlyLysIleMetCysTyrValGlyAspValArgSerVal	467
Db	1948	GCTGTGGTGAACCGGTGAAGCAGAGGGGGAATATCAATACGTGAACGACGTGAGGAACATC	2007
Qy	468	ThrGlnLysHisIleGlnGluTyrGlyProPheAspLeuValIleGluGlySerProCys	487
Db	2008	ACAAAGAAATATATTGAAGATGGGGCCCATTTGACTTGGTATGGCGGAAAGCCCATGC	2067
Qy	488	AsnAspLeuSerIleValAsnProAlaArgLysGlyLeuTyrGlnGlyThrGlyArgLeu	507
Db	2068	AACAGATCTCTCAATGTGAATCCAGCCCGAGAAAGGCGTGTATGAGGTACAGCCGGCTC	2127
Qy	508	PhePheGluPheTyrArgGlyLeuLeuHisAspAlaArgProLysGluLysAspArgPro	527
Db	2128	TTCTTCGAATTATTCACCTCTGTATATCTACGCCCCAAGAGAGGAGATGACCGGGCGG	2187
Qy	528	PhePheTrpLeuPheGluAsnValValAlaMetGlyValSerAspLysArgAspIleSer	547
Db	2188	TTCTTCTCGAATCTTGAAGATGTTGTACCCATGAAGCTTTGGCGCAAGAGGAGCATCTCA	2247
Qy	548	ArgPheLeuGluSerAsnProValMetIleAspAlaLysGluValSerAlaAlaHisArg	567

Db	2248	CGGTTCTCGGAGGTGTATCCAGTCGATGATTCAGTCCATCAAAAGTTTCTGCTGCACAGG	2307
Oy	568	AlaArgTyrPheThrPGlyAsnLeuProGlyMetAsnArgProLeuAlaSerThrValAsn	587
Db	2308	GCCCGAATCTTCGGGGCAACTTACCCGGAGTGAACAGGCCCGATGATGATCAAGAT	2367
Oy	588	AspLysLeuGluLeuGlnGluCysLeuGluLysGluLysArgIleAlaLysPheSerLysVal	607
Db	2368	GATAAATCTCGAGCTGCAGAGACTGCTTGGAATACAAATGAGATAGCCAAAGTTAAAGAA	2427
Oy	608	ArgThrIleLeuThrThrArgSerAsnSerIleLysGlnGlyLysAspGlnHisPheProVal	627
Db	2428	CAGCAATTAACACCAAGTGAATCTCATATAACAGGGGAAAAACCAACTTTTCCCTGTT	2487
Oy	628	PheMetArgLysGluAspIleLeuTyrCysThrGluMetGluArgValPheGlyPhe	647
Db	2488	GTCATGAATGGCAAGAAGATGTTTGTGGTGCATGACTCGAAAGATCTTTGGCTTT	2547
Oy	648	ProValHisTyrThrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeuGly	667
Db	2548	CTGTGCACTACACAGACGTGTCCACATGAGCGCGTGTGTCGCCCGCAGAGCTGCTGGGA	2607
Oy	668	ArgSerTyrSerValProValIleArgHisLeuPheAlaProLeuLysGluTyrPheAla	687
Db	2608	AGGCTCTGGAGCGTGCCTGTCAATCCAGACCTTTCGCCCCCTCTGAAGGACTACTTGCA	2667
Oy	688	Cys 688	
Db	2668	TGT 2670	
RESULT 15			
US-10-144-577-38			
; Sequence 38, Application US/10144577			
; Publication No. US20030083292A1			
; GENERAL INFORMATION:			
; APPLICANT: Macleod, Alan Robert			
; TITLE OF INVENTION: Inhibitors of DNA Methyltransferase Isoforms			
; FILE REFERENCE: MET-005			
; CURRENT APPLICATION NUMBER: US/10/144,577			
; PRIOR FILING DATE: 2002-05-13			
; PRIOR APPLICATION NUMBER: US 60/290,202			
; PRIOR FILING DATE: 2001-05-11			
; PRIOR APPLICATION NUMBER: US 60/290,212			
; PRIOR FILING DATE: 2001-05-11			
; NUMBER OF SEQ ID NOS: 49			
; SOFTWARE: FastSeq For Windows Version 4.0			
; SEQ ID NO 38			
; LENGTH: 4145			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-144-577-38			
Alignment Scores:			
Pred. No.:			
Score: 1.09e-234			
Percent Similarity: 2159.50			
Best local Similarity: 71.8%			
Query Match: 58.5%			
DB: 57.5%			
Gaps: 7			
US-10-623-813-85 (1-689) x US-10-144-577-38 (1-4145)			
Oy	10	SeqGlySerGlnLysValGluGluAlaSerProProAlaValGlnGlnProThrAsp	29
Db	623	TCACAGCGACACAGAGGACACACACGTGGACGCCCCAGACAGAGATACCCCTACG---	679
Oy	30	ProLaserProThrValAlaThr-----ThrProGluProValGlyGlyAsp	45
Db	680	CCCGCCTAGCCCGACGACGCCACAGCGGGGGGCATGTGATCCCGCAG-CTGGAGGCCAGAC	738
Oy	46	AlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGluTyrGluAspGlyArg	65
Db	739	AGTGGAGAT-----GGAGACAGTTCAAGATGATCAGATGGGAG	777

QY 66 GlyPheGlyIleGlyLeuValTTPGlyLysLeuArgIlyPheSerTrpTrpProGly 85
 DB 778 GAGTTTGGAAATAGGAGCCCTCGTGTGGGAAAGATCAAGGGCTTCTCTGGGGCCGCC 837
 QY 86 ArgIleValSerTrpTrpMetThrGlyArgSerArgAlaIleGlyIleThrArgTTPVal 105
 DB 838 ATGGTGTGTCTTGGAAAGGCCACCTCCAGCCAGAGCTTGTCTGTGGCATGGCGTGGTTC 897
 QY 106 MetTrpPheGlyAspGlyLysPheSerValValCysValGlyLysLeuMetProLeuSer 125
 DB 898 CAGTGTGTTGGCATGCGAAGTTCTCCGAGGTCTTCGACAGCAAACTGGGGGCACTGGGG 957
 QY 126 SerPheCysSerAlaPheHisGlnAlaATrTyrAsnLysGlnProMetTyrArgLysAla 145
 DB 958 CTGTTACGCCACCACTTAAATTGGCCACCTTCATAAAGCTCGCTTCCTATCGAAAGCC 1017
 QY 146 IleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeuPheProAlaCysHis 165
 DB 1018 ATGTTACATGCTCTGGAGAAAGCTTAGGGTGGAGCTGGCAAGACTTCCCC----- 1068
 QY 166 AspSerAspGluSerAspSerGlyLysAlaValAlaGluValGlnAsnLysGlnMetIleGlu 185
 DB 1069 -----AGCAGCCCTGGAGACTATTGGAGCAGACCACTGAAGCCCATGTTGGAG 1116
 QY 186 TTPAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLysLeuLysProGluGlu---- 204
 DB 1117 TGGGCCACAGGGGGCTTCAAGCCCACTGGGATCGAGGGCTTCAAAACCAACACAGCAA 1176
 QY 205 -----GluLysAsnPro 208
 DB 1177 CCAGTGTAAATTAAGTCGAAGGTGCTCGTCGACGAGCAATGGAATTAATCAAGAAA 1236
 QY 209 TyrLysGluValTyrThrAspMetTrpValGluProGluAlaAla---TyrAlaPro 227
 DB 1237 TACGAGAAACAGACTCGAAGACGACAGCTGACGACTGACGCCCTGACTGACTGCCCC 1296
 QY 228 ProProProAlaLysLysProArgLysSerThrThrGluLysProLysValLysGluIle 247
 DB 1297 GCACCCCAAGCGCTTCAACAAATTGCTTAACACAGCCGAAAGACGGGGATGAA--- 1353
 QY 248 IleAspGluArgThrArgGluArgLeuValTyrGluValArgLysCysArgAsnIle 267
 DB 1354 -----GATCAGAGCCGAGAACAAATGGCTTCAGATGTTGCCAACACAGACCACTCG 1407
 QY 268 GluAspIleCysIleSerCysGlySerLeuAsnValThrLeuGlnHisProLeuPheIle 287
 DB 1408 GAAGATGGCTGTTGTCTTGTGGCAGAAACCCCGTGTCTTCCACCTCTCTTGAAG 1467
 QY 288 GlyGlyMetCysGlnAsnLysLysAsnLysPheLeuGluCysAlaTyrGlnTyrAspAsp 307
 DB 1468 GGGGGGCTCTGTCAAGCATGCCGGGATCCCTTCTTGAAGCTGTTTACATGATGTAGC 1527
 QY 308 AspGlyTyrGlnSerTyrCysThrIleCysCysGlyIleArgGluValLeuMetCysGly 327
 DB 1528 GATGGCTATCACTTACTGCACTGTGTCTGTGGAGGGCCGAGAGCTGCTGCTTTCAGC 1587
 QY 328 AsnAsnAsnCysCysArgCysPheCysValGluCysValAspLeuLeuValGlyProGly 347
 DB 1588 AACAGAGAGCTGCTGGCGGTGTTCTGTGTGAGAGTGCCTGAGAGGTGCTGGGGCAGAGC 1647
 QY 348 AlaAlaGlnAlaAlaIleLysGluAspProTTPAsnLysTyrMetCysGlyHisLysGly 367
 DB 1648 ACAGCGGCGGAGCCCAAGCTTCAAGAGCCCTGAGCTGTACATGTGTCTCCGCGAGCCG 1707
 QY 368 ThrTyrGlyLeuLeuArgArgArgGluAspTrpProSerArgLeuGlnMetPhePheAla 387
 DB 1708 TGTCAATGGCTCTGGGGCGCGGAGAGCATGTGAAGTGGCTGCAAGGCTTCTTCAAGC 1767
 QY 388 AsnAsnHisAspGlnGluPheAspProProLysValTyrTrpProValProAlaGluLys 407
 DB 1768 AGTGAACAGGGGCTTGAATATACAAAGCCCAAGCTGTACCTGCATTCGCCGAGCCGGA 1827

QY 408 ArgLysProIleArgValLeuSerLeuPheAspGlyIleAlaThrGlyLeuLeuValLeu 427
 DB 1828 AGCGGAGCCCATTCAGAGCTCTGCTCATTTGTTATGGCATGCGACAGCTCACTAGCTTC 1887
 QY 428 LysAspLeuGlyIleGlnValAspArgTyrIleAlaSerGluValCysGluAspSerIle 447
 DB 1888 AAGAGTGGGCATTAAGGTAGGAAGTACGCTGCTTGAAGTGTGAGAGATGCATT 1947
 QY 448 ThrValGlyMetValArgHisGlnGlyLysIleMetTyrValGlyAspValArgSerVal 467
 DB 1948 GCTGTGGAAACCGTAAGCAGCAGGGGGAATATCAATAGCTAAGCAGCTGAGGAACATC 2007
 QY 468 ThrGlnLysHisIleGlnGluTrpGlyProPheAspLeuValIleGlyGlySerProCys 487
 DB 2008 ACMAAGAAAAATATGAAGAAAGGGGCCATTGTGACTTGATGTATGGCGGAAGCCATGC 2067
 QY 488 AsnAspLeuSerIleValAsnProAlaArgLysGlyLysLeuTyrGluGlyThrGlyArgLeu 507
 DB 2068 AACGATCTCTCAATGTGAATCCAGCCAGAAAGCCCTGTATGAGGGTACAGGCCGCTC 2127
 QY 508 PhePheGluPheTyrArgLeuLeuHisAspAlaArgProLysGluGlyAspAspArgPro 527
 DB 2128 TTCTTCGAATTTTACCTCCTGTAATTAATCAAGCCCAAGAGGCTGATGACCGGCG 2187
 QY 528 PhePheTrpLeuPheGluAsnValAlaIleMetGlyValSerAspLysArgAspIleSer 547
 DB 2188 TTCTTCTGATTTTGAGAAATTTGTAGCCATGAAAGTTGGGGGACAGAGGAGCATCTCA 2247
 QY 548 ArgPheLeuGluSerAsnProValIleIleAspAlaLysGluValSerAlaHisArg 567
 DB 2248 CGGTTCTGAGTGAATTCAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 2307
 QY 568 AlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArgProLeuAlaSerThrValAsn 587
 DB 2308 GCCCATCTCTTGGGGCAACTACCCGGAGAGAACAGGCCGTGTATGATCAAAAGAT 2367
 QY 588 AspLysLeuGluLeuGlnGluCysLeuGlnHisGlyArgIleAlaLysPheSerLysVal 607
 DB 2368 GATTAATCTGAGCTCAGAGCTGTTGAAATACATAGATAGCCAAAGTTAAAGAAATGA 2427
 QY 608 ArgThrIleThrThrArgSerAsnSerIleLysGlnGlyLysAspGlnHisPheProVal 627
 DB 2428 CAGCAATATACACACAAAGTCGAATCGATCAACAGGGGAAAAACCACTTTCCTGTT 2487
 QY 628 PheMetAsnGluLysGluAspIleLeuTrpCysThrGlnMetGluArgValPheGlyPhe 647
 DB 2488 GTCATGAATGGCAAGAAAGATGTTTGTGTGCTGACCTGAGCTGAAAGGATCTTGTGCTTT 2547
 QY 648 ProValHisTyrThrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeuGly 667
 DB 2548 CTTGTGCACTACACAGACGTGTCCAACTGGGCGCTGTGCGCCGCAAGAGCTGTGGGA 2607
 QY 668 ArgSerTrpSerValProValIleArgHisLeuPheAlaProLeuLysGluTyrPheAla 687
 DB 2608 AGCTCCTGAGGCTCCTGTATCCGACACCTTTCGCCCCCTGTGAAGGATCACTTGTGA 2667
 QY 688 Cys 688
 DB 2668 TGT 2670

Search completed: October 10, 2006, 15:22:14
 Job time : 2142 secs